



Eunice Cristina dos Reis Beijinho

Licenciada em Engenharia Informática

Modelos de risco para inimigos do Brócolo

Dissertação para obtenção do Grau de Mestre em
Engenharia Informática

Orientador: Carlos Augusto Isaac Piló Viegas Damásio, Professor
Associado, Universidade Nova de Lisboa
Co-orientador: João Carlos Gomes Moura Pires, Professor Auxiliar,
Universidade Nova de Lisboa

Júri

Presidente: Doutora Ana Maria Diniz Moreira
Vogais: Doutor Luís Miguel Mendonça Rato
Doutor Carlos Augusto Isaac Piló Viegas Damásio



FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Setembro, 2018

Modelos de risco para inimigos do Brócolo

Copyright © Eunice Cristina dos Reis Beijinho, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa.

A Faculdade de Ciências e Tecnologia e a Universidade NOVA de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

AGRADECIMENTOS

Começo por agradecer ao meu orientador, Carlos Viegas Damásio, pela oportunidade de trabalhar num projeto com potencial para fazer a diferença e pelo apoio e direção que me deu durante o desenvolvimento dos modelos e escrita desta dissertação.

Quero agradecer também ao meu co-orientador João Carlos Gomes Moura Pires pelo apoio e ensino de metodologias de trabalho importantes.

Agradeço também à Faculdade de Ciências e Tecnologia, em especial aos professores e professoras do Departamento de Informática, por me terem acolhido e me terem permitido adquirir os conhecimentos necessários para construir um futuro na área da engenharia informática e crescer profissionalmente.

Não posso deixar de agradecer aos meus colegas de projeto por construírem os outros componentes desta solução, em especial ao (agora) mestre Francisco Pinto.

Por fim, quero agradecer à minha família pelo tempo e carinho que sempre estiveram dispostos a dar, pelo apoio e pelos vários conselhos que levo para a minha vida adulta.

RESUMO

A produção de brócolo em Portugal tem sofridos danos de grande dimensão, causando o abandono da produção por parte de alguns empresários agrícolas. Os prejuízos económicos associados ao cultivo do brócolo devem-se maioritariamente aos inimigos *Alternaria brassicicola* (fungo) e *Delia radicum* (mosca da couve), que são capazes de destruir hectares de terreno agrícola num curto espaço de tempo. É necessário compreender melhor estes organismos para podermos combatê-los, construindo e validando modelos de desenvolvimento dos mesmos para o território português. Esta dissertação vem apresentar um sistema que se propõe a validar e adaptar os modelos de previsão destes inimigos utilizando dados meteorológicos históricos assim como informação recolhida pelos agricultores. Através destes modelos, é possível construir cartas de risco passíveis de serem utilizadas pelos empresários agrícolas por forma a garantir um maior controlo sobre as suas culturas e um indicador da melhor altura para aplicar os produtos fitossanitários.

Palavras-chave: *Alternaria brassicicola*, *Delia radicum*, cartas de risco, Portugal

ABSTRACT

The production of broccoli in Portugal has suffered great damages, causing the withdrawal of its continuation by some agricultural entrepreneurs. The economic losses associated with broccoli cultivation are mainly due to the enemies *Alternaria brassicicola* (fungus) and *Delia radicum* (cabbage fly), which are able to destroy hectares of agricultural land in a very short time period. It is necessary to better understand these organisms so that we can fight them. Development models should be constructed and validated for the portuguese territory. This dissertation presents a project that proposes to validate and adapt the existing predictive models of these enemies using past weather data and information provided by the farmers and, with these models, construct risk maps that can be used by the agricultural entrepreneurs in order to guarantee a greater control over their cultures and provide an indicator of the best time to apply the phytosanitary products.

Keywords: *Alternaria brassicicola*, *Delia radicum*, risk maps, Portugal

ÍNDICE

Lista de Figuras	xiii
Lista de Tabelas	xv
Listagens	xvii
Glossário	xix
1 Introdução	1
1.1 Enquadramento e Motivação	1
1.2 Delia radicum	2
1.2.1 Ciclo de vida	3
1.3 Alternariose do brócolo	3
1.3.1 Ciclo de vida	4
1.4 Problema	5
1.5 Objetivos e contributos	5
1.6 Estrutura do documento	6
2 Trabalho relacionado	7
2.1 Variáveis meteorológicas relevantes para modelos de pragas e doenças . .	7
2.2 Modelos de desenvolvimento para Alternaria	9
2.2.1 Modelo genérico de infeção para patologias fúngicas foliares . . .	9
2.2.2 Modelo HDC FV53b	17
2.3 Modelos de desenvolvimento para Delia radicum	18
2.3.1 Métodos para cálculo de graus-dia	18
2.3.2 Modelos de desenvolvimento da mosca da couve	21
2.4 Obtenção do valor de folha molhada	22
2.4.1 Limiar de humidade relativa constante	22
2.4.2 CART/SLD	23
2.4.3 Limiar de DPD constante	24
2.4.4 Limiar de VPD constante	25
2.5 Conclusões	25
3 Solução	27

3.1	Visão global do sistema	27
3.2	Arquitetura e componentes da solução	28
3.2.1	Leitura e tratamento dos dados	28
3.2.2	Implementação e execução dos modelos de risco	32
3.2.3	Construção das cartas de risco	34
3.2.4	Integração das cartas de risco na plataforma	35
3.3	Modelo de dados	37
3.3.1	Relação estacoes	38
3.3.2	Relação estacao_mais_proxima	39
3.3.3	Relação series_temporais	39
3.3.4	Relação modelo_serie_temporais	39
3.3.5	Relações parametros_series_temporais e nomes_parametros . .	40
3.3.6	Relação tipodados	40
3.3.7	Relação alertas	40
3.4	Tecnologias utilizadas	41
3.5	Desafios encontrados e soluções	41
3.5.1	Leitura e tratamentos dos dados	42
3.5.2	Implementação dos modelos de risco	42
3.5.3	Construção das cartas de risco	42
3.5.4	Conclusão	43
4	Resultados obtidos	45
4.0.1	Modelos de folha molhada	45
4.1	Validação dos modelos	46
4.1.1	Modelo 5 horas de folha molhada - M5HM	46
4.1.2	Modelo HDC FV53b	48
4.1.3	Modelo genérico de infecção para patologias fúngicas foliares - soma - MGIPFF-S	49
4.1.4	Modelo genérico de infecção para patologias fúngicas foliares - per- centagem - MGIPFF-P	51
4.1.5	Conclusões	52
4.2	Modelo Delia Radicum	53
4.3	Desempenho da solução	54
4.4	Conclusões	54
5	Conclusões e trabalho futuro	55
5.1	Conclusões	55
5.2	Trabalho futuro	55
	Bibliografia	57

LISTA DE FIGURAS

1.1	<i>Delia radicum</i> [5]	2
1.2	Ciclo de vida da mosca da couve (<i>Delia radicum</i>). Origem das imagens: [19] .	3
1.3	<i>Alternaria brassicicola</i> (x650) [25]	4
1.4	Brócolo afetado por alternariose [2]	4
1.5	Ciclo de vida de <i>Alternaria brassicicola</i> [9]	5
2.1	Função de resposta à temperatura T	10
2.2	Função W(T)	10
2.3	Modelo genérico de infeção para patologias fúngicas foliares - utilização horária sugerida por Simone Bregaglio (adaptado para português)	12
2.4	Modelo genérico de infeção para patologias fúngicas foliares - utilização horária através da soma de horas molhadas	14
2.5	Modelo genérico de infeção para patologias fúngicas foliares - utilização horária através da soma de horas molhadas	16
2.6	Modelo HDC FV53b - Condições para ocorrência de infeção	17
2.7	Modelo HDC FV53b - Condições para ocorrência de esporulação	17
2.8	Método do seno para calcular o número de graus-dia [22]	19
2.9	Método do seno duplo para calcular o número de graus-dia [11]	20
2.10	Método do triângulo simples para calcular o número de graus-dia [11]	20
2.11	Método do triângulo duplo para calcular o número de graus-dia [11]	20
2.12	Ciclo de vida da <i>Delia radicum</i> utilizado pelo projeto SafeBrócolo[6]	22
2.13	Modelo estendido do limiar de humidade relativa (Kruit, Roy J. Wichink, et al,2004)	23
2.14	Modelo CART/SLD - Kim et al.(2002)	24
3.1	Visão global do sistema	28
3.2	Estações meteorológicas do IPMA no território continental português	29
3.3	Estações meteorológicas do IPMA no arquipélago dos Açores	29
3.4	Estações meteorológicas do IPMA na Região Autónoma da Madeira	29
3.5	Carta de risco de alternariose apresentada pela plataforma	36
3.6	Página com o risco associado a uma parcela	37
3.7	Página que permite visualizar o histórico do risco de todas as estações meteorológicas	37

3.8	Modelo de dados da solução	38
4.1	Risco de alternariose segundo o modelo M5HM para Outubro de 2015	47
4.2	Risco de alternariose segundo o modelo HDC FV53b, comparação entre poli- nómios	49
4.3	Risco de alternariose segundo o modelo HDC FV53b com o máximo dos poli- nómios para Outubro de 2015	50
4.4	Risco de alternariose segundo o modelo MGIPFF-S	51
4.5	Risco de alternariose segundo o modelo MGIPFF-P para Outubro de 2015 . .	52
4.6	Modelo de desenvolvimento de <i>Delia radicum</i> - Projeto SafeBrócolo	53

LISTA DE TABELAS

2.1	Exemplo da aplicação do método da média para cálculo dos graus-dia	19
3.1	Dados meteorológicos antes da interpolação	30
3.2	Dados meteorológicos depois da interpolação	31
3.3	Exemplo dos dados presentes no mapa de valores de risco	34
3.4	Escala de risco de alternaria	34
3.5	Escala de risco de ataque por mosca da couve	35
4.1	Matriz de confusão para estimativa de folha molhada a partir de déficit de pressão de vapor (DPV) para a estação da Chamusca (setembro 2015 a janeiro de 2017)	46
4.2	Funções encontradas por interpolação empírica dos dados conhecidos para o modelo HDC FV53b	48

LISTAGENS

4.1	Algoritmo do modelo M5HM	47
4.2	Algoritmo do modelo HDC FV53b	48

GLOSSÁRIO

API	<i>Application Programming Interface</i> - é um conjunto de rotinas e padrões estabelecidos por um software para a utilização das suas funcionalidades por aplicações que não pretendem envolver-se em detalhes da implementação do software, mas apenas usar os seus serviços..
DataFrame	Estrutura de dados usada pela biblioteca pandas. Assemelha-se a uma tabela, com linhas e colunas. Para mais informação, visitar https://pandas.pydata.org/pandas-docs/stable/generated/pandas.DataFrame.html .
endpoint	URL associado a um serviço fornecido por uma API .
geography	Tipo de dados que representa um ponto espacial com coordenadas geográficas.
JSON	JavaScript Object Notation é um formato open-source de troca de dados simples e rápida entre sistemas. Para mais informação, visitar http://json.org/ .
parcela	Terreno com finalidade agrícola.

INTRODUÇÃO

Neste capítulo é descrita a motivação para a elaboração desta dissertação e o projeto onde se enquadra. São apresentados os dois maiores inimigos do brócolo em Portugal e é explicado o problema em questão. São enumerados os contributos desta dissertação e efetuado um pequeno resumo da estrutura do documento.

1.1 Enquadramento e Motivação

O desenvolvimento da agricultura foi dos passos mais importantes que a humanidade deu durante a sua existência. Passámos a produzir o nosso alimento em vez de ter de o perseguir, e prosperámos como espécie. Com o passar do tempo, começámos a depender cada vez mais desta atividade, o que trouxe algumas desvantagens: as plantas estão sujeitas a condições meteorológicas desfavoráveis, pragas e outras doenças, tornando-as uma fonte de alimento frágil. Com o avanço da ciência, foram sendo descobertas formas de combater estas fragilidades, exemplos incluem as estufas e os produtos fitossanitários.

Para que seja possível uma aplicação mais responsável e eficaz destes produtos é necessário um estudo da cultura e dos inimigos a combater. Em condições ideais, os pesticidas seriam aplicados na data em que o seu impacto fosse maior, em quantidades adequadas; em vez disso, os produtos fitossanitários são aplicados muitas vezes em datas fixas, sem conhecimento do risco real dos inimigos associados.

Em Portugal, o grupo “SafeBrócolo - Melhoria do Processo Produtivo com base em modelos de risco para alternaria e mosca da couve” estudou os dois maiores inimigos do brócolo: *Alternaria brassicicola* (fungo) e *Delia radicum* (mosca da couve). O grupo foi constituído por membros de diversas instituições, entre elas:

- “FCT/UNL — Faculdade Ciências e Tecnologia, Universidade Nova de Lisboa”;
- “UE — Universidade de Évora”;

- “ISA — Instituto Superior de Agronomia”;
- “ESAS — Escola Superior Agrária de Santarém”;
- “FNOP — Federação Nacional de Organizações de Produtores”;
- “CADOVA — Cooperativa Agrícola Do Vale De Arraiolos”.
- “COTHN — Centro Operativo e Tecnológico Hortofrutícola Nacional”;
- “INIAV — Instituto Nacional de Investigação Agrária e Veterinária”;

O seu principal objetivo foi a construção de cartas de risco relativas a Portugal para estes dois inimigos, suportadas por modelos de previsão.

Estas cartas de risco servem como importante auxílio aos empresários agrícolas, permitindo uma melhor gestão e agendamento dos produtos fitossanitários. O desenvolvimento deste projeto foi possível graças ao financiamento do programa “PRODER - Programa de Desenvolvimento Rural”, que existe no âmbito da estratégia “Europa 2020”.

1.2 *Delia radicum*



Figura 1.1: *Delia radicum* [5]

Popularmente designada como mosca da couve, *Delia radicum* (figura 1.1) é conhecida por estragos de grande dimensão e custo nas plantações de Brassicas (couves, nabos, brócolo, nabiça, etc.) [20]. Esta praga pode ter até 4 gerações por ano e está mais ativa entre Abril e Outubro [3, p. 38].

Sintomas comuns em plantas infestadas:

- Atraso no crescimento
- Larvas nas raízes

- Murchidão
- Caule destruído

1.2.1 Ciclo de vida

Este inseto passa por quatro fases durante a sua vida (figura 1.2):

1. Ovo - A primeira fase da vida da mosca da couve. Os ovos (com cerca de 1mm) são depositados perto dos caules das plantas, em pequenos aglomerados e assim ficam durante cerca de uma semana (4 a 6 dias). Para que o desenvolvimento ocorra, é necessária uma temperatura do ar entre os 15°C e 20°C [8].
2. Larva - Quando os ovos eclodem, as larvas saem e enterram-se, alimentando-se das partes menos duras da raiz da planta. É nesta fase que a mosca da couve causa mais estragos. O desenvolvimento continua durante mais 2-3 semanas [1, p. 26], até que as larvas se transformam em pupas.
3. Pupa - A mosca da couve passa cerca de 2 semanas nesta fase. Este período pode prolongar-se caso a temperatura do solo esteja demasiado alta (acima de 22°C nos meses de verão) ou demasiado baixa (nos meses de inverno) [8]. Nesse caso, as pupas entram em diapausa até que a temperatura volte a um nível aceitável.
4. Adulto - Depois de sair da pupa e se tornar adulta, a mosca da couve reproduz-se nos primeiros 5 dias [1, p. 25] e começa a depositar os ovos a partir do 3º dia, depositando mais de 100 ovos [8].

1.3 Alternariose do brócolo

A alternariose é uma doença fúngica provocada pelos organismos da espécie *Alternaria* como *Alternaria brassicae* e *Alternaria brassicicola*. Estes inimigos atacam principalmente

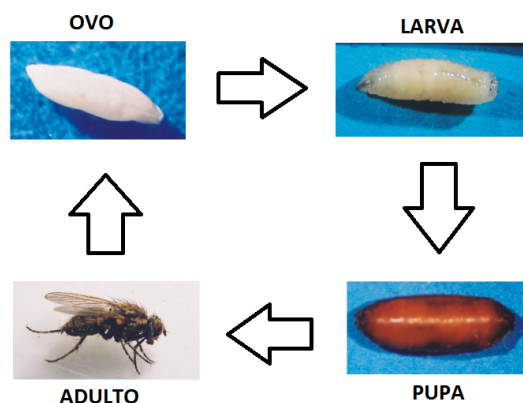


Figura 1.2: Ciclo de vida da mosca da couve (*Delia radicum*). Origem das imagens: [19]

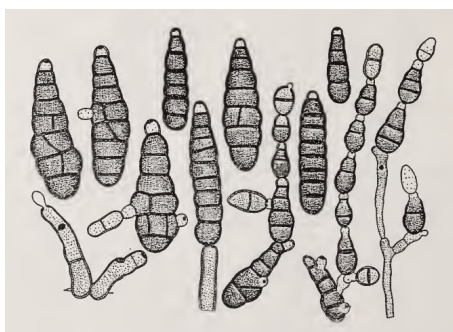


Figura 1.3: *Alternaria brassicicola* (x650) [25]

plantas da família *Brassica* (nabo, couve-flor, brócolo, etc.) e levam a grandes perdas e prejuízo por parte dos empresários agrícolas [25, p. 4]. O fungo desta família mais danoso para as culturas de brócolo é a *Alternaria brassicicola* [2] (figura 1.3), mais ativo em Portugal entre os meses de Setembro e Dezembro.



Figura 1.4: Brócolo afetado por alternariose [2]

Sintomas de brócolos infetados: [25, p. 9]

- Cabeça acastanhada (figura 1.4)
- Lesões escuras no eixo principal
- Manchas escuras e circulares nas folhas
- Sementes infetadas são mais pequenas, apresentam descoloração e por vezes fungos visíveis.

1.3.1 Ciclo de vida

O ciclo de vida deste fungo (observável na figura 1.5) é o seguinte:

1. A infeção ocorre quando o fungo está presente nas folhas ou vagens e existem condições favoráveis. As condições favoráveis para infeção são temperaturas entre os 15 e 25°C e um período de folha molhada entre 4 a 5 horas. A infeção pode ocorrer fora deste intervalo de temperaturas, desde que a folha se mantenha molhada por mais tempo.

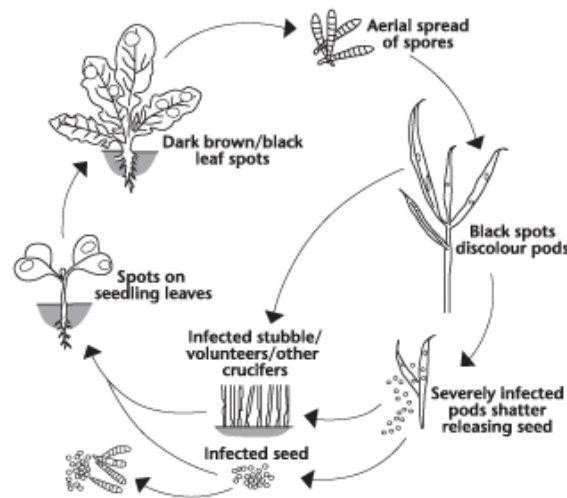


Figura 1.5: Ciclo de vida de *Alternaria brassicicola* [9]

2. Depois da planta infetada, começam a formar-se mitosporos. Na presença de uma humidade relativa alta e/ou períodos de folha molhada, a esporulação aumenta, causando os sintomas acima descritos. Estes sintomas surgem 4 a 6 dias depois da infeção.
3. Quando existe uma queda na humidade relativa, os esporos são libertados e transportados pelo vento e pela chuva até às plantas vizinhas, recomeçando o ciclo.

Por vezes, as próprias sementes já estão infetadas com esta doença. Resíduos de plantas infetadas de culturas anteriores são também uma fonte de infeção.

1.4 Problema

Alternaria brassicicola e *Delia radicum* são dois organismos influenciados por fatores climáticos como a temperatura do ar, a precipitação e até o momento da rega[2], sendo que existem várias subpopulações de cada um deles, cada uma com características diferentes dependendo do país de origem e do clima do mesmo. Estes inimigos foram estudados para subpopulações de outros países e foram desenvolvidos modelos de previsão, mas os modelos de previsão gerados relativos a outros países podem não se adequar a Portugal, dada a diferença entre climas e subpopulações. É necessário ajustar e validar estes modelos de forma a aumentar a sua previsibilidade para o território português.

1.5 Objetivos e contributos

Antes de mais, foi feito um estado da arte de modelos de desenvolvimento e previsão de *Alternaria brassicicola* e *Delia radicum*. Foram avaliados esses mesmos modelos com dados reais meteorológicos e de presença de *Alternaria brassicicola* e *Delia radicum* históricos. Depois dos modelos avaliados, foi implementado um conjunto de modelos (com parâmetros

para o território português) e um sistema de apoio para elaboração de cartas de risco para os inimigos especificados e para a cultura do brócolo, tendo como ponto de partida dados meteorológicos de estações locais (colocadas nos campos agrícolas), estações do "IPMA - Instituto Português do Mar e da Atmosfera", e dados de satélite. Por fim, foi feita a integração com uma plataforma pública de apoio aos produtores de brócolo desenvolvida pelo projeto (<http://safebrocolo.fct.unl.pt/>), permitindo a visualização das cartas de risco na plataforma, bem como um histórico dos índices de risco para cada terreno agrícola. Os modelos de previsão desenvolvidos são específicos para as subpopulações de Portugal. As cartas de risco permitem aos agricultores portugueses um maior controlo sobre as suas culturas do brócolo e servem como aviso antes dos estragos terem sido feitos.

1.6 Estrutura do documento

Neste primeiro capítulo foi introduzido o problema da ausência de modelos de previsão para Portugal e a motivação para o seu desenvolvimento e implementação. Foi descrito brevemente o projeto SafeBrócolo e foram enumerados os contributos esperados.

No capítulo seguinte são aprofundados os dois inimigos sobre os quais o projeto se debruçou: *Alternaria brassicicola* e *Delia radicum*. São explicados os modelos de previsão e desenvolvimento existentes e as variáveis meteorológicas relevantes para os mesmos. Por fim, são enumerados alguns métodos de obtenção do valor de folha molhada, um dos parâmetros mais importantes para os modelos e apresentadas algumas conclusões.

O capítulo 3 descreve ao pormenor a solução desenvolvida, a sua arquitetura, os componentes e os passos envolvidos no seu funcionamento. É apresentado o modelo de dados que foi criado para dar suporte à solução, as tecnologias utilizadas e por fim os desafios que surgiram durante o desenvolvimento e as soluções encontradas para os resolver.

No capítulo 4 são apresentados os resultados obtidos, sendo apresentadas as conclusões sobre o melhor modelo de folha molhada, avaliados os modelos de alternariose e da mosca da couve, e por fim apresentados os resultados sobre o desempenho temporal e espacial da solução desenvolvida.

Finalmente, no capítulo 5 é feita uma conclusão sobre o projeto e são discutidas ideias de possíveis trabalhos futuros.

TRABALHO RELACIONADO

Foram pesquisados modelos de desenvolvimento e previsão para os inimigos *Alternaria brassicicola* e *Delia radicum*. Mais especificamente, os modelos de desenvolvimento para a *Alternaria* foram encontrados no Web of Science, através da pesquisa "prediction alternaria" e "Alternaria brassic*", limitadas pela categoria "AGRICULTURE MULTIDISCIPLINARY". Foram também consultados os parceiros do projeto SafeBrócolo, que indicaram alguns modelos de desenvolvimento para a mosca da couve.

O capítulo começa por incluir uma pequena descrição das variáveis de entrada normalmente presentes em modelos de previsão de doenças agrícolas. De seguida, são explicados os modelos de previsão encontrados e que podem ser adaptados para os inimigos estudados. Finalmente, são enumerados os métodos de obtenção do valor da folha molhada e apresentadas algumas conclusões.

2.1 Variáveis meteorológicas relevantes para modelos de pragas e doenças

Os parâmetros mais utilizados na modelação de pragas e doenças agrícolas são:

- T_{air} - Temperatura do ar. Esta variável é importante pois pode ser relacionada com as temperaturas cardeais dos organismos e usada para estimar o ritmo de desenvolvimento dos mesmos[26]. Neste documento as unidades desta variável serão os graus Celsius.
- RH - Humidade relativa (0-100%). Representa a percentagem de vapor de água presente no ar, em relação à quantidade máxima que o ar consegue reter à temperatura atual. Este parâmetro é importante para determinar se as folhas estão molhadas, condição relacionada com infeções fúngicas[18][10].

- Temperatura do solo - Esta temperatura varia não só com a localização mas também com a profundidade à qual a medição foi realizada e o tipo de solo. Neste documento as unidades desta variável serão os graus Celsius.
- *VPD* - Défice de pressão de vapor (Vapor Pressure Deficit). Representa a diferença entre a quantidade de vapor de água presente no ar (VP_{air}) e a quantidade de vapor de água que o ar consegue suportar, quando saturado (VP_{sat}). Este parâmetro pode ser calculado através da equação apresentada por Prenger e Ling em [23, p. 6]:

$$VPD = vp_{sat} - vp_{air}$$

em que

$$vp_{sat} = e^{(-1.044 \times 10^4/T) + (-1.129 \times 10^1) + (-2.702 \times 10^{-2} \times T) + (1.289 \times 10^{-5} \times T^2) + (-2.478 \times 10^{-9} \times T^3) + (6.456 \times \ln(T))}$$

$$vp_{air} = vp_{sat} \times RH/100 \text{ e } T = (T_{air} + 273.15) \times 9/5$$

O resultado desta equação resulta num valor em psi. Para utilizar a unidade SI Pascal, basta multiplicar o resultado por 6.89475729 e obtém-se o VPD em kilo Pascal (kPa). Estudos mostram que um VPD abaixo de 0.43 kPa é favorável à sobrevivência de patógenos fúngicos e que a infeção é maior abaixo dos 0.20 kPa [23, p. 3].

- *DPD* - Diferença para o ponto de orvalho (Dew Point Depression). Representa a diferença entre a temperatura do ar e a temperatura do ponto de orvalho. Esta informação pode também ser utilizada para determinar se as folhas estão molhadas ou não [15]. O valor *DPD* pode ser calculado pela fórmula

$$DPD = T_{air} - T_{dew}$$

em que T_{dew} representa a temperatura do ponto de orvalho. Esta última pode ser calculada através da equação apresentada por Mark G. Lawrence em 2005 [17]:

$$T_{dew} = \frac{B \left[\ln\left(\frac{RH}{100}\right) + \frac{A \times T_{air}}{B + T_{air}} \right]}{A - \ln\left(\frac{RH}{100}\right) - \frac{A \times T_{air}}{B + T_{air}}}$$

em que $A = 17.625$ e $B = 243.04$. Neste documento as unidades desta variável serão os graus Celsius.

- *WS* - Velocidade do vento (Wind Speed). Neste documento as unidades desta variável serão os km/h. Existem modelos que, para determinar se as folhas de uma planta estiveram molhadas, utilizam esta variável [15].
- Duração do período de folha molhada - Tão ou mais importante que a verificação da presença de água nas folhas, é a duração em horas dessa presença. Vários estudos e modelos de infeção utilizam esta duração para estimar a incidência da doença [18][10].
- Precipitação - Medida em mm. Tem um papel importante na dispersão dos esporos de inimigos fúngicos e contribui para a presença de água nas folhas.

2.2 Modelos de desenvolvimento para Alternaria

Neste capítulo serão apresentados dois modelos de patologias fúngicas promissores encontrados na literatura: o modelo genérico de infecção para patologias fúngicas foliares e o modelo HDC FV53b, que é mais específico para o gênero *Alternaria*.

2.2.1 Modelo genérico de infecção para patologias fúngicas foliares

O modelo desenvolvido por R. D. Magarey, T. B. Sutton, and C. L. Thayer [18] permite calcular a duração em horas do período de folha molhada necessário para atingir 20% incidência ou 5% severidade de doença a uma dada temperatura T . Para produzir este valor, o modelo usa:

- T - a temperatura média durante o período de folha molhada (graus Celsius);
- T_{min} - a temperatura mínima para que a infecção ocorra (graus Celsius);
- T_{max} - a temperatura máxima à qual a infecção pode ocorrer (graus Celsius);
- T_{opt} - a temperatura ótima para infecção (graus Celsius);
- W_{min} - período de folha molhada mínimo para atingir um nível crítico de infecção a qualquer temperatura (em horas);
- W_{max} - a duração ótima do período de folha molhada para que ocorra infecção
- D_{50} - máxima duração do período seco, em horas. Um valor superior resulta numa diminuição de 50% na infecção.

As temperaturas cardeais de infecção do fungo (T_{min} , T_{max} e T_{opt}) são usadas para calcular a resposta à temperatura atual. A função de resposta à temperatura usada é baseada num modelo que descreve o desenvolvimento de culturas com base na temperatura (Yin et. al [26]).

$$f(T) = \left(\frac{T_{max} - T}{T_{max} - T_{opt}} \right) \left(\frac{T - T_{min}}{T_{opt} - T_{min}} \right)^{(T_{opt} - T_{min}) / (T_{max} - T_{opt})}$$

Como exemplo, pode ser observada na figura 2.1 a variação do ritmo de desenvolvimento para o fungo *Alternaria brassicae*, usando os valores cardeais $T_{min}=2.6^{\circ}\text{C}$, $T_{opt}=18^{\circ}\text{C}$ e $T_{max}=35^{\circ}\text{C}$ presentes na literatura[18, p. 4].

A infecção tem um valor nulo de resposta à temperatura nos valores cardeais limite (T_{min} e T_{max}) e atinge o valor máximo de resposta à temperatura(1) aos 18°C , a temperatura ótima para infecção.

Este modelo utiliza esta função para calcular W_T - a duração em horas do período de folha molhada necessário para atingir 20% incidência ou 5% severidade de doença a uma dada temperatura T .

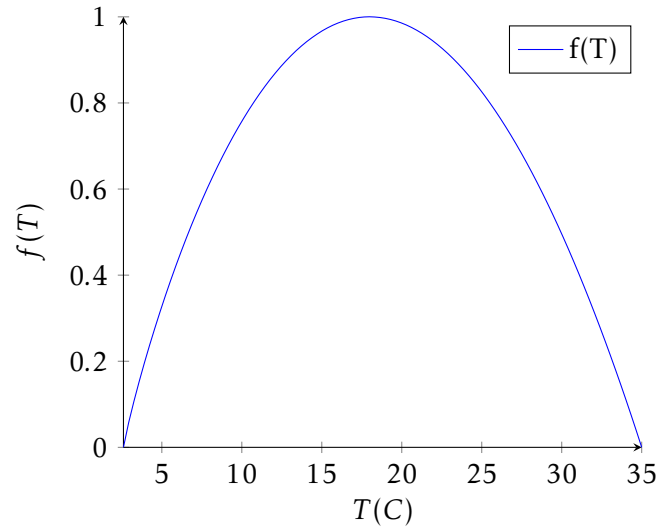


Figura 2.1: Função de resposta à temperatura T

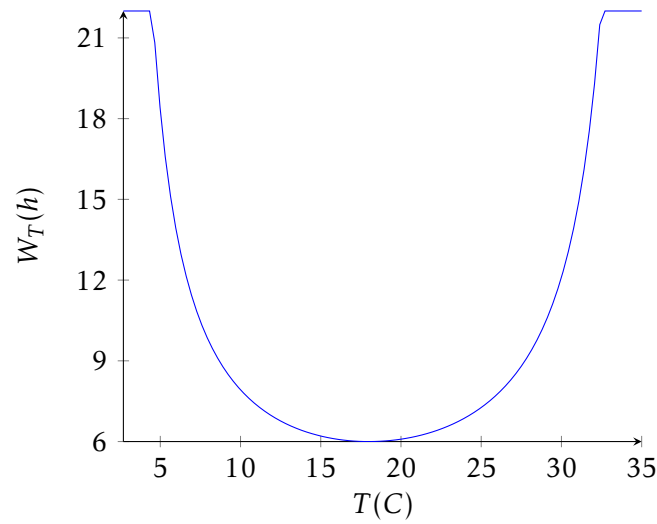


Figura 2.2: Função W(T)

$$W_T = \begin{cases} \frac{W_{min}}{f(T)} & \text{se } \frac{W_{min}}{f(T)} \leq W_{max} \\ W_{max} & \text{caso contrário} \end{cases}$$

Utilizando a função de resposta à temperatura $f(T)$ anteriormente descrita e os valores $W_{min} = 6$ horas e $W_{max} = 22$ horas correspondentes ao fungo *Alternaria brassicae*¹ [18, p. 4] obtém-se (figura 2.2):

Como seria de esperar, as temperaturas mais próximas da temperatura ótima para infecção (18°C) correspondem aos valores mínimos desta função e o valor máximo é 22 horas (W_{max}).

¹O artigo não providencia os valores correspondentes a *Alternaria brassicicola*, portanto este exemplo é relativo a *Alternaria brassicae*.

O parâmetro D_{50} é utilizado no caso em que os dados meteorológicos são recebidos de hora em hora em vez de diariamente. Se D_{50} horas seguidas de período seco acontecerem, a infecção diminui 50%.

Uma das vantagens deste modelo é que todos os parâmetros de entrada têm um significado biológico claro. Este modelo ignora a relação entre a temperatura e a humidade relativa ainda assim mas consegue prever corretamente o período de folha molhada mínimo para níveis críticos de infecção para uma variedade de patógenos fúngicos [18, p. 4].

Contudo, Magarey, Sutton e Thayer não aprofundam a metodologia de utilização do modelo no caso de serem recebidos dados meteorológicos horários.

Simone Bregaglio sugere, na sua dissertação de doutoramento, um funcionamento deste modelo [4, pp. 70–71] com dados horários (figura 2.3). T_h representa a temperatura de cada hora.

Em cada hora, a temperatura atual é comparada com as temperaturas cardeais de infecção. Se esta for válida, é criada uma nova coorte de esporos. De seguida, todas as coortes são iteradas para verificar se cumprem os requisitos para a ocorrência de infecção. Se for o caso, é lançado um evento de potencial infecção.

Seja T a temperatura média do período de folha molhada, os requisitos para que seja lançado o alerta de potencial infecção para uma dada hora e coorte são:

- A folha esteve molhada na hora presente;
- A temperatura atual está dentro dos limites cardeais de infecção;
- A duração do período de folha molhada necessária para atingir o nível crítico de infecção W_T para a temperatura média do período de folha molhada T está entre $W_{min}/f(T)$ e W_{max} .

A terceira condição desta lista não é correta. Se a temperatura média do período de folha molhada estiver dentro dos limites cardeais, então a condição será sempre cumprida (ver função W_T acima), resultando no lançamento de alertas mesmo quando W_T assume um valor alto, ou seja de várias horas. Na realidade, quer-se que sejam lançados eventos de potencial infecção apenas passadas as W_T horas.

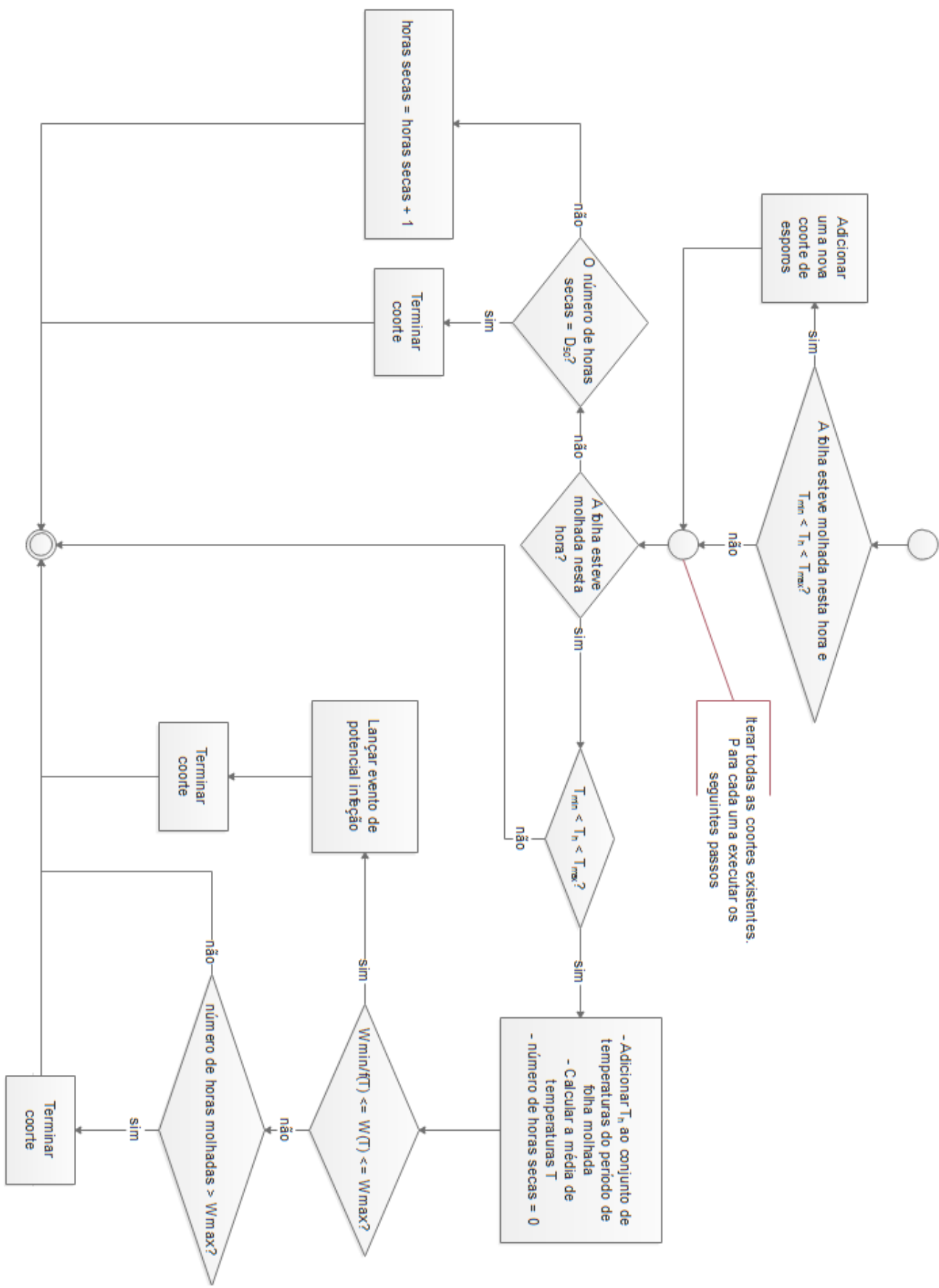


Figura 2.3: Modelo genérico de infecção para patologias fúngicas foliares - utilização horária sugerida por Simone Bregaglio (adaptado para português)

Como alternativa, sugere-se o modelo presente na figura 2.4 que compara o número de horas propícias à infecção (chamadas horas molhadas) já passadas com as necessárias para atingir o nível crítico de infecção (W_T).

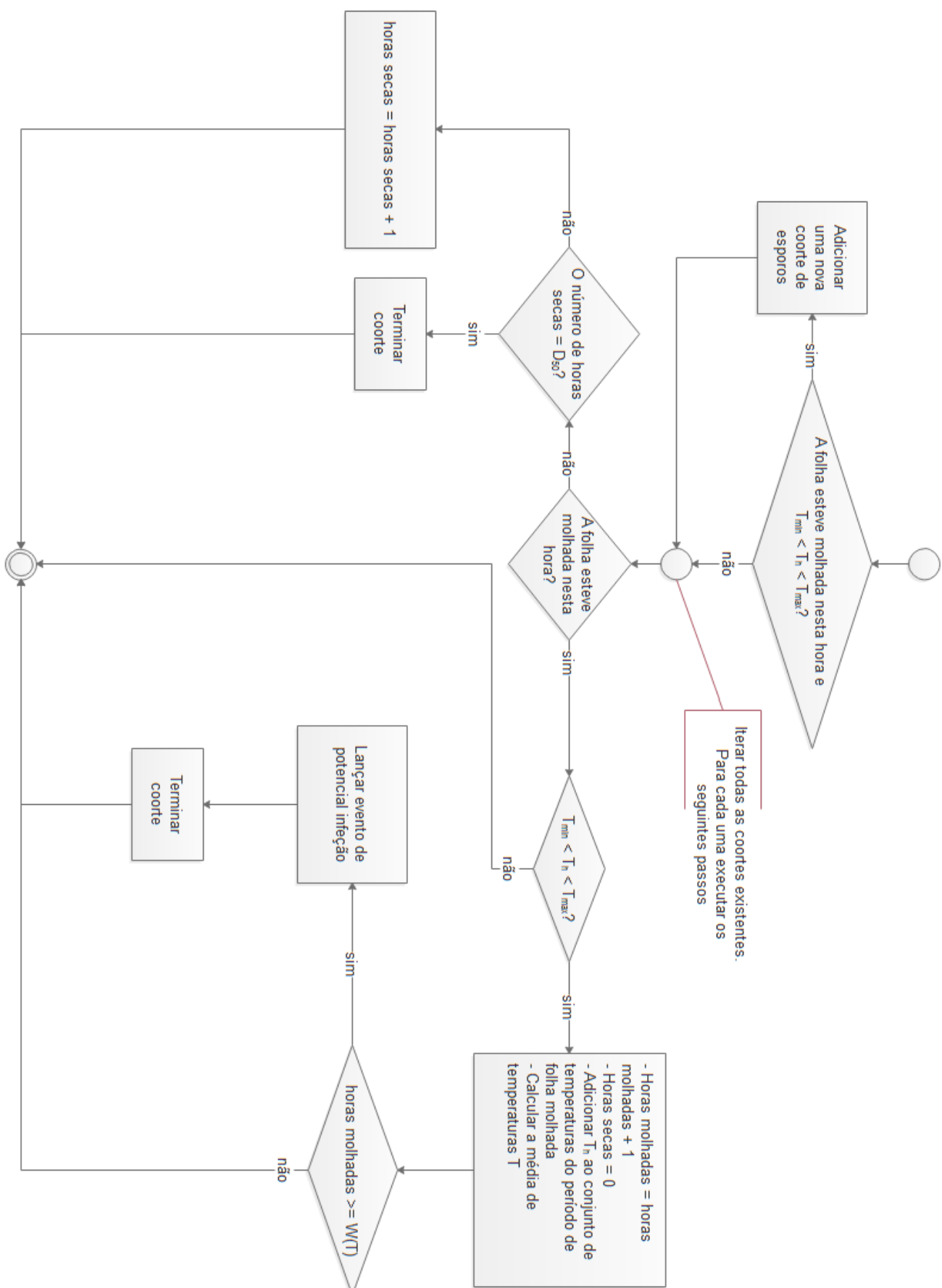


Figura 2.4: Modelo genérico de infecção para patologias fúngicas foliares - utilização horária através da soma de horas molhadas

Vale a pena mencionar também uma segunda alternativa que coordena o lançamento de alertas com base numa variável DI (Desenvolvimento Infecção com imagem 0-1). DI começa com valor 0 para todas as coortes e, em cada hora propícia, aumenta em $1/W(T_h)$. $1/W(T_h)$ representa a parte que uma hora (a uma dada temperatura T_h) contribuiu para o tempo total necessário. Quando a variável assume o valor 1, então foi atingido o nível crítico da doença e é emitido o alerta. O diagrama deste modelo pode ser consultado na figura 2.5.

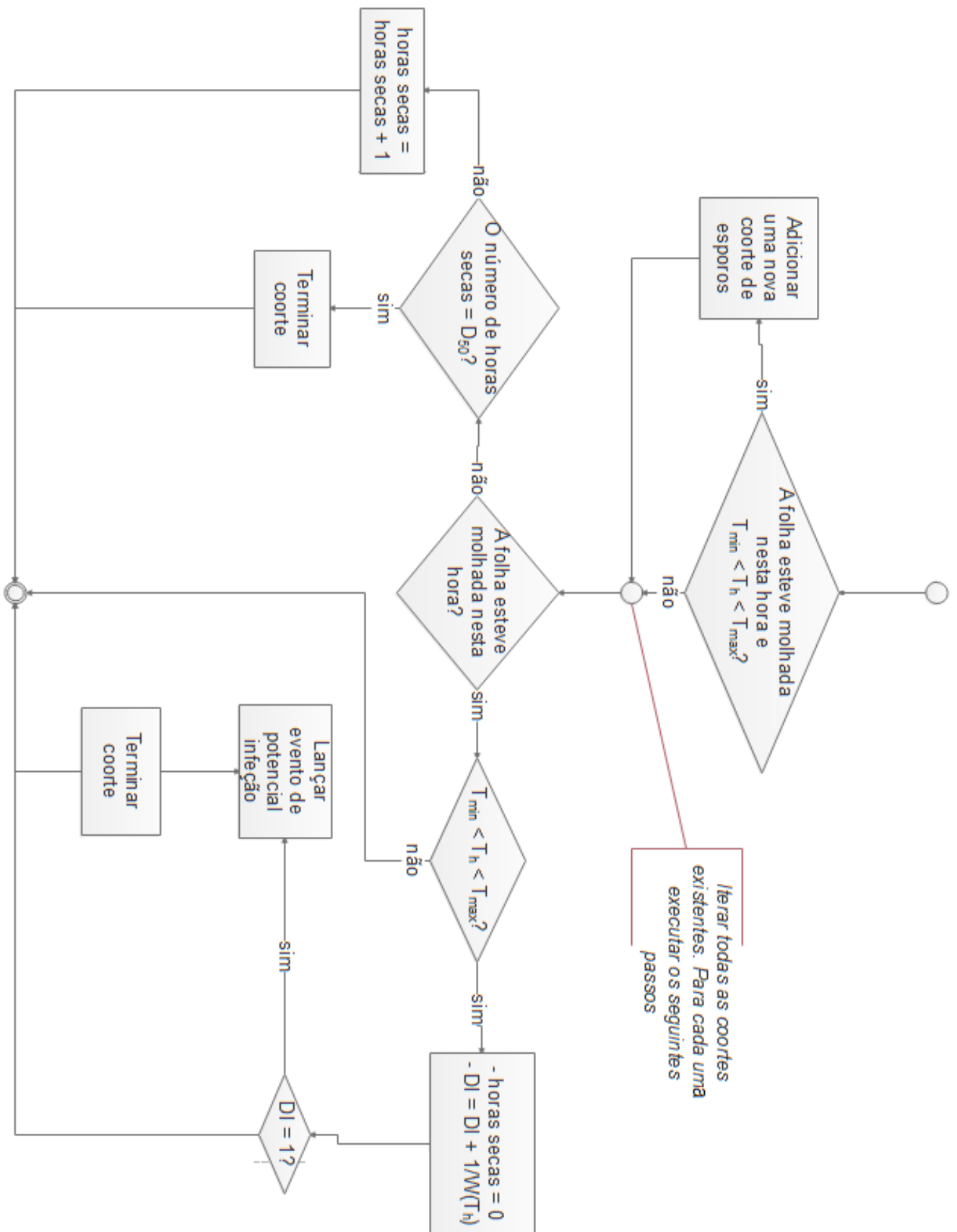


Figura 2.5: Modelo genérico de infecção para patologias fúngicas foliares - utilização horária através da soma de horas molhadas

2.2.2 Modelo HDC FV53b

O modelo desenvolvido por R. Kennedy e A.M. Graham [13] permite determinar se existe infecção ou esporulação com base na duração do período de folha molhada ou no número de horas em que existiu uma alta humidade. Os autores efetuaram um estudo sobre as condições meteorológicas associadas à infecção e esporulação do agente patogénico *Pyrenopeziza brassicae*, um dos inimigos das plantas brássicas como o brócolo.

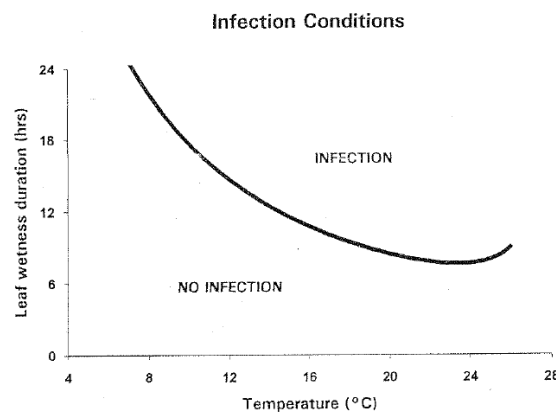


Figura 2.6: Modelo HDC FV53b - Condições para ocorrência de infecção

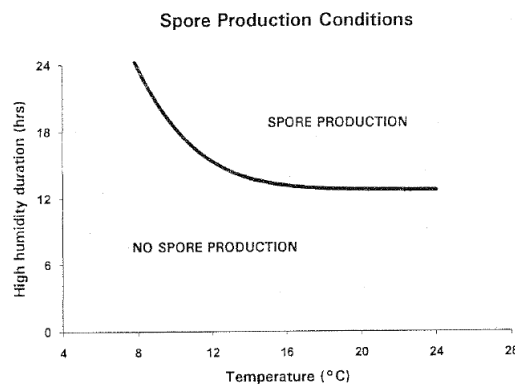


Figura 2.7: Modelo HDC FV53b - Condições para ocorrência de esporulação

O modelo apresenta dois gráficos, apresentados nas figuras 2.6 e 2.7. O primeiro gráfico determina as condições para que seja considerada uma potencial infecção e o segundo as condições para a ocorrência de esporulação. Por exemplo, um período de folha molhada de 18 horas a uma temperatura média de 16°C resultaria em infecção, segundo este modelo.

Os autores não disponibilizam as funções representadas nos gráficos mas pode ser feita a interpolação empírica através dos pontos conhecidos.

2.3 Modelos de desenvolvimento para *Delia radicum*

Os insetos são animais de sangue frio, logo a temperatura à sua volta tem um impacto maior do que nos animais de sangue quente. Na verdade, todos os insetos precisam de uma dada acumulação de calor para atingir cada uma das suas fases de vida. Esta acumulação de calor é medida em graus-dia.

Um grau-dia corresponde a um grau acumulado num dia completo (24 horas) em que a temperatura do ar se manteve entre os limites cardiais de desenvolvimento do inseto. Para que seja possível o cálculo do número de graus-dia são necessários quatro parâmetros: o limite inferior de desenvolvimento do inseto (em °C), o limite superior de desenvolvimento (em °C) e as temperaturas máxima e mínima do dia.

Nem todos os insetos têm um limite superior de desenvolvimento. Quando o limite superior de desenvolvimento é atingido, o desenvolvimento do inseto pode tornar-se constante ou até mesmo parar. Cada fase de desenvolvimento de um inseto tem as suas necessidades de acumulação de calor.

2.3.1 Métodos para cálculo de graus-dia

Existem várias maneiras de calcular o número de graus-dia acumulados:

2.3.1.1 Método da média

Os graus-dia podem ser calculados usando apenas a temperatura média diária. Sendo:

- T_{min} a temperatura mínima do dia
- T_{max} a temperatura máxima do dia
- T a temperatura média do dia calculada através da expressão $[T_{max} + T_{min}]/2$.
- LI o limite inferior de desenvolvimento
- LS o limite superior de desenvolvimento

em cada dia são acumulados $T - LI$ graus-dia, se $LI < T < LS$. Como exemplo, imaginemos um inseto com uma temperatura mínima de desenvolvimento de 10°C e que necessita de 8 graus-dia para passar à fase de desenvolvimento seguinte. No cenário em que T representa a temperatura média do dia, é possível determinar o dia em que o inseto atinge a fase de desenvolvimento seguinte (ver tabela 2.1). No quinto dia, o inseto atingiu a fase de desenvolvimento seguinte.

2.3.1.2 Método da média modificado

O problema do método da média surge quando $T_{min} < LI$ ou $T_{max} > LS$. Se a temperatura mínima de um dia ocorreu apenas durante 1 hora, o número de graus dia acumulados

Tabela 2.1: Exemplo da aplicação do método da média para cálculo dos graus-dia

Dia	T	Graus-dia	Graus-dia acumulados
1	12°C	2	2
2	14°C	4	6
3	9°C	0	6
4	10°C	0	6
5	12°C	2	8

pode estar errado visto que o método anterior apenas tem em conta a temperatura média do dia. Para corrigir este problema, o método da média modificado usa a seguinte expressão:

$$T = \begin{cases} [T_{max} + T_{min}]/2 & \text{se } T_{min} > LI \text{ e } T_{max} < LS \\ [T_{max} + LI]/2 & \text{se } T_{min} < LI \text{ e } T_{max} < LS \\ [LS + T_{min}]/2 & \text{se } T_{min} > LI \text{ e } T_{max} > LS \\ [LS + LI]/2 & \text{se } T_{min} < LI \text{ e } T_{max} > LS \end{cases}$$

2.3.1.3 Método do seno

O método do seno baseia-se na assunção de que a temperatura durante um dado dia tem um comportamento sinusoidal. Para calcular o número de graus-dia usando este método, é calculada a área da curva dentro dos limites cardeais do inseto (figura 2.8).

O método do seno simples utiliza a regressão do seno e desenha uma curva simétrica em volta da temperatura máxima de cada dia.

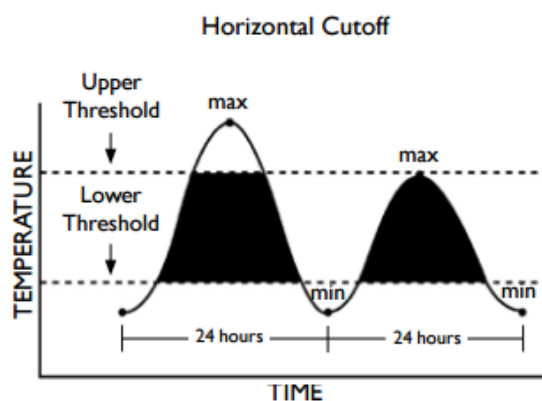


Figura 2.8: Método do seno para calcular o número de graus-dia [22]

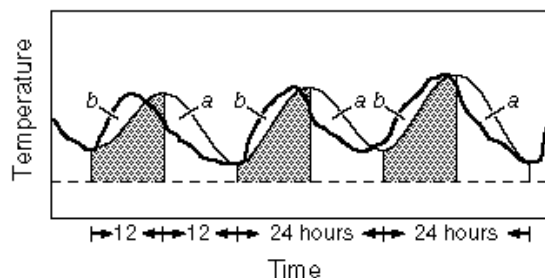


Figura 2.9: Método do seno duplo para calcular o número de graus-dia [11]

O método do seno duplo desenha, para cada dia, duas curvas. A primeira curva desde a temperatura mínima do dia anterior até à temperatura máxima do próprio dia e a segunda da temperatura máxima até à mínima do dia seguinte (figura 2.9).

2.3.1.4 Método do triângulo

O método do triângulo traça uma linha reta entre a temperatura mínima do dia e a máxima e vice-versa. Como as temperaturas estão limitadas, muitas vezes o resultado é, para cada dia, um triângulo. Para calcular o número de graus dia é calculada a área de cada triângulo dentro dos limites cardeais do organismo (figura 2.10).

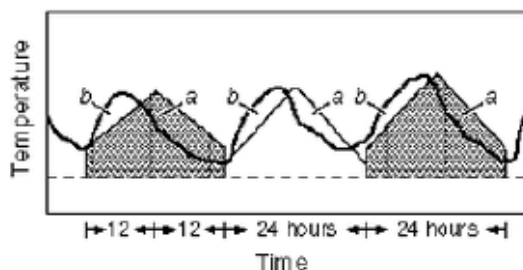


Figura 2.10: Método do triângulo simples para calcular o número de graus-dia [11]

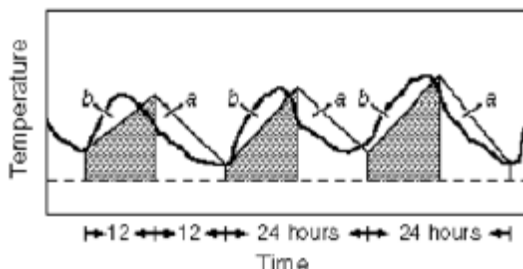


Figura 2.11: Método do triângulo duplo para calcular o número de graus-dia [11]

Uma variante deste método é o método do triângulo duplo, em que as temperaturas são medidas de 12 em 12h. Os graus-dia totais são obtidos através da soma das áreas dos dois lados do triângulo (figura 2.11).

2.3.2 Modelos de desenvolvimento da mosca da couve

Modelo de Eckenrode e Chapman, 1971 Este modelo sugere que a primeira emergência de adultos surge quando acumulados 166.7 graus-dia e que uma geração de *Delia radicum* (de adulto a adulto) dura 653.3 graus-dia. Afirma também que o limite inferior de desenvolvimento LI é 6.1°C. O estudo foi feito em Wisconsin (EUA).

Modelo de Eckenrode e Chapman, 1972 Este modelo surge um ano depois do primeiro, pelos mesmos autores. Foi usado o mesmo limite inferior de desenvolvimento $LI = 6.1^\circ\text{C}$. Desta vez, a primeira aparição de adultos surge quando acumulados 135 graus-dia e uma geração de *Delia radicum* (de adulto a adulto) dura 617.8 graus-dia. O estudo foi feito no mesmo local (Wisconsin, EUA).

Modelo de Vincent e Stewart, 1981 Este modelo usa um limite inferior de desenvolvimento ligeiramente mais baixo (6.0°C). Sugere que a primeira aparição de adultos surge quando acumulados 192 graus-dia e atinge um máximo entre os 314 e 398 graus-dia. O segundo pico de adultos surge entre os 847 graus-dia e os 960 graus-dia e o terceiro pico entre os 1446 e os 1604 graus-dia. O estudo foi realizado na província do Quebec, no Canadá.

Modelo de Bažok, Ceranić-Sertić, Barčić et al., 2012 O modelo desenvolvido em 2012 na Croácia mostra um valor de 746 graus-dia entre gerações de *Delia radicum* e sugere que o primeiro pico de adultos surge quando acumulados 471.35 ± 74.97 graus-dia. O segundo pico surge aos $1,217.28 \pm 96.12$ graus-dia. Para este estudo foi usado um limite inferior de desenvolvimento de 4.3°C.

Modelo SafeBrócolo I, 2017 Os investigadores do projeto SafeBrócolo desenvolveram também um modelo de desenvolvimento da mosca da couve *Delia radicum*). O limite inferior de desenvolvimento usado foi 6°C e o primeiro e segundo pico de adultos surgem aos 130 e 727 graus-dia, respetivamente. Este modelo sugere uma duração de 580 graus-dia para cada geração. Na figura 2.12 é possível observar o ciclo de vida da *Delia radicum* sugerido por este modelo.

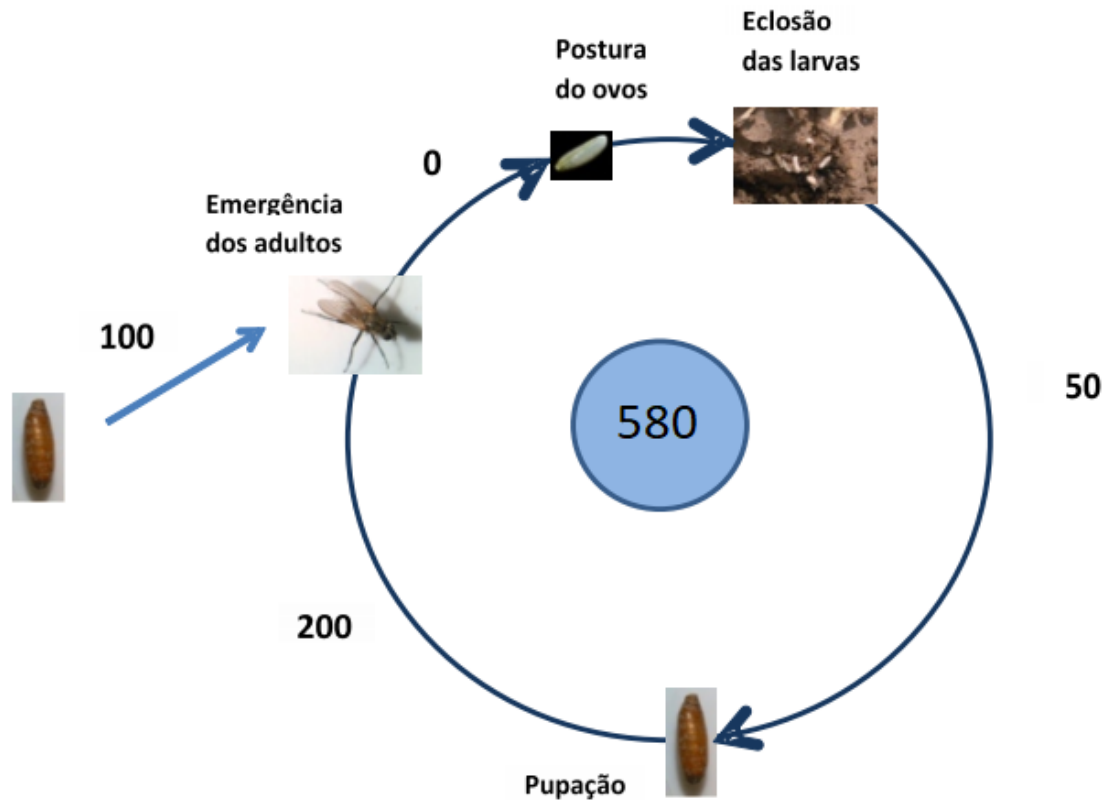


Figura 2.12: Ciclo de vida da *Delia radicum* utilizado pelo projeto SafeBrócolo[6]

2.4 Obtenção do valor de folha molhada

A obtenção do valor de folha molhada LW (sim/não) em cada hora não está disponível através de dados meteorológicos fornecidos pelas estações meteorológicas públicas ou por satélite. Existem estações capazes de detetar a folha molhada, mas são propriedade privada de empresas ou particulares. Os dados fornecidos pelas estações meteorológicas do Instituto Português do Mar e da Atmosfera (IPMA) não incluem, por enquanto, informação acerca da folha molhada.

Assim sendo, é necessário utilizar modelos científicos para estimar o valor deste parâmetro.

2.4.1 Limiar de humidade relativa constante

Este modelo sugere que as folhas estão molhadas se a humidade relativa for maior que um predeterminado limiar. Esta assunção baseia-se em estudos que mostram que a condensação nas folhas começa antes do ar ficar saturado, quando a humidade relativa está entre os 91 e 99% [21]. Por exemplo, se o limiar definido for de 91% então todas as horas em que a humidade relativa for maior ou igual a 91%, são consideradas horas molhadas.

$$RH \geq 91 \Rightarrow \text{hora molhada}$$

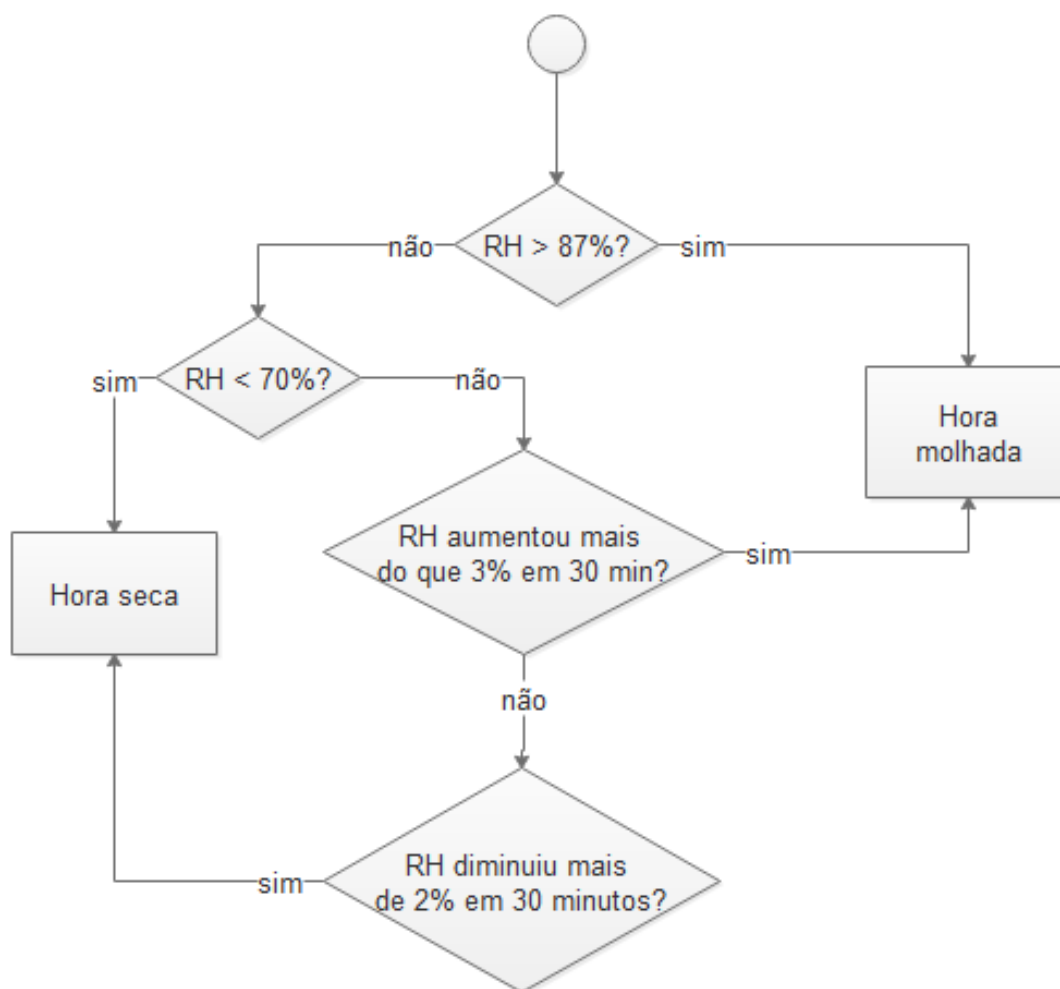


Figura 2.13: Modelo estendido do limiar de humidade relativa (Kruit, Roy J. Wichink, et al, 2004)

Um modelo um pouco mais complexo baseado neste é o modelo estendido do limiar de humidade relativa (figura 2.13), proposto em 2004 por Kruit, Roy J. Wichink, et al. [16], que diz que num dado período de tempo em que a humidade relativa varie entre os 70% e os 87%, a hora é considerada molhada se a humidade relativa aumentar mais do que 3% em meia-hora. Se a humidade relativa diminuir mais do que 2% em meia-hora, então a hora é considerada seca. Horas em que a humidade relativa RH seja maior que 87% ou menor que 70% são automaticamente consideradas molhadas ou secas, respetivamente.

2.4.2 CART/SLD

O método CART/SLD proposto em 2002 por Kim et al. baseia-se na aplicação da árvore de classificação apresentada na figura 2.14. Em primeiro lugar é verificado se a diferença de temperaturas entre a temperatura média durante a hora e a temperatura de ponto de orvalho (Dew Point Depression ou DPD) é menor que 3.7°C. Caso a diferença seja maior que 3.7°C, a hora é considerada seca, ou seja, não existiu ocorrência de folha molhada.

Caso o valor de DPD seja menor que 3.7°C, é verificada a velocidade do vento. Se esta for menor que 9km/h, a hora é considerada molhada se a inequação A for verdadeira.

Inequação A:

$$(1.6064 \times \sqrt{T}) + (0.0036 \times T^2) + (0.1531 \times RH) - (0.4599 \times WS \times DPD) - (0.0035 \times T \times RH) > 14.4674$$

em que T é a temperatura média em graus Celsius, WS é a velocidade do vento em m/s, RH é a humidade relativa(0-100%) e DPD é a diferença de temperaturas entre T e a temperatura de ponto de orvalho.

Se a velocidade do vento for maior que 9km/h, é humidade relativa é comparada com 87.8%. Se for menor, a hora é considerada seca. Se for maior, a hora é molhada se a inequação B for cumprida.

Inequação B:

$$(0.7921 \times \sqrt{T}) + (0.0046 \times RH^2) - (2.3889 \times WS) - (0.039 \times T \times WS) + (1.0613 \times WS \times DPD) > 37$$

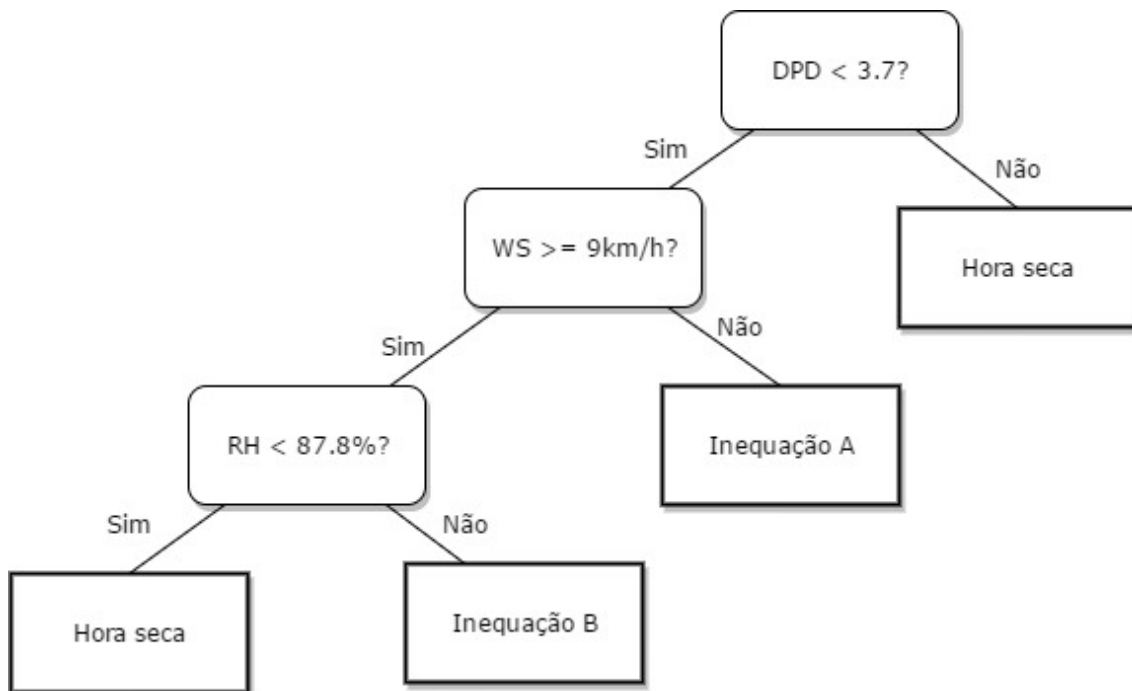


Figura 2.14: Modelo CART/SLD - Kim et al.(2002)

2.4.3 Limiar de DPD constante

Alguns autores[12] sugerem o uso da variável DPD (ver 2.1) para estimar a presença de folha molhada. O modelo é simples: Se numa dada hora o DPD for menor que um limiar predefinido, então a hora é molhada. Por exemplo, se o limiar for 2°C:

$$DPD < 2^{\circ}\text{C} \Rightarrow \text{hora molhada}$$

2.4.4 Limiar de VPD constante

Visto que o VPD está intimamente relacionado com evapotranspiração, este pode ser usado para verificar a existência de gotículas de água no ar, assumindo que um valor menor que 3 kPa indica que a folha está molhada[14]:

$$VPD \leq 3kPa \Rightarrow \text{hora molhada}$$

2.5 Conclusões

Os modelos de infecção fúngica apresentados serão implementados, validados e comparados durante o decorrer deste projeto, de modo a obter o maior número de previsões corretas.

Os modelos de desenvolvimento da mosca da couve foram estudados pelo projeto SafeBrócolo e por opção do mesmo, o implementado foi o modelo SafeBrócolo I.

Em relação aos vários métodos de obtenção dos graus-dia, estudos feitos anteriormente pela ex-aluna da FCT-UNL Filipa Ferreira mostram que o método do seno duplo é o mais adequado para a *Delia radicum* [7] no território português.

Simone Bregaglio[4] sugere que, dos modelos para a estimativa do valor de folha molhada apresentados, o modelo CART é o mais apropriado. Esta conclusão pode não ser válida para o território português, uma vez que os dados meteorológicos utilizados pelo autor para o estudo dos vários modelos de folha molhada são relativos a outros países. Por esta razão, serão também validados os vários modelos de folha molhada.

SOLUÇÃO

Neste capítulo será descrita e estudada a solução adotada para o problema apresentado neste documento. Será primeiramente apresentado o processo de escolha dos modelos de risco cujos valores são disponibilizados na plataforma online. Posteriormente, será apresentada a arquitetura do sistema desenvolvido bem como o modelo de dados que foi construído como extensão ao já utilizado pela base de dados da plataforma pública para dar suporte à integração das cartas de risco. Serão listadas as tecnologias utilizadas e apresentados alguns detalhes importantes sobre a implementação de um sistema desta natureza.

3.1 Visão global do sistema

A figura 3.1 mostra a visão global do sistema desenvolvido. O sistema está a executar continuamente numa máquina com o sistema operativo Ubuntu 16.04.3 LTS e todos os dias às 01h recolhe os novos dados disponibilizados pelo IPMA. O diagrama mostra a relação entre os vários componentes que constituem a solução desenvolvida.

Os dados do IPMA são recolhidos pelo componente IPMAConnector. São depois passados ao componente InfoGetter que é responsável por tratar os dados e preencher dados em falta.

Os dados já tratados são então enviados aos componentes responsáveis por executar os modelos (RiskUpdater ou DCCorrector). Estes componentes utilizam o DBConnector para armazenar os valores de risco calculados na base de dados do projeto, construída sobre PostgreSQL 9.5.10.

Depois de executados os modelos, o componente produtor de rasters gera as cartas de risco e disponibiliza-as para a plataforma do projeto SafeBrócolo (disponível no endereço <http://safebrocolo.di.fct.unl.pt/>).

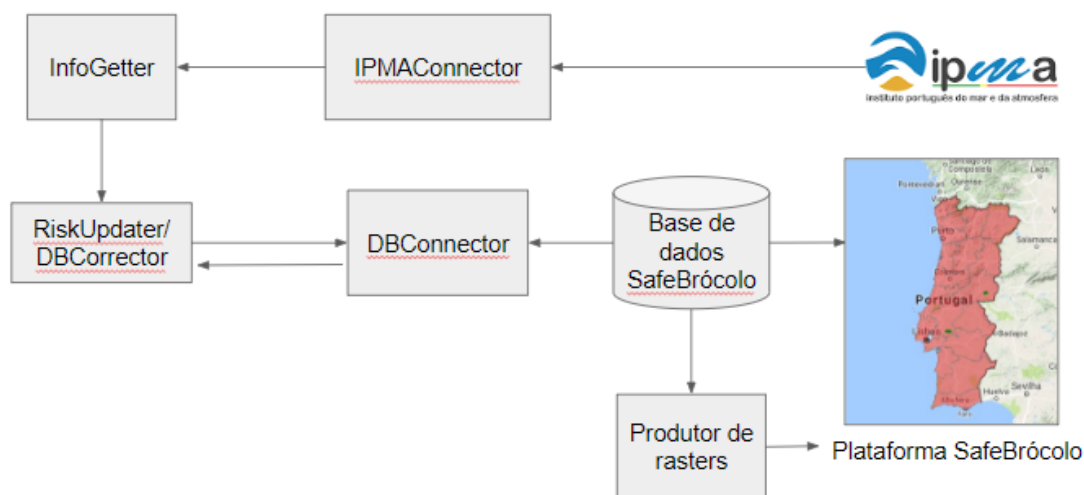


Figura 3.1: Visão global do sistema

3.2 Arquitetura e componentes da solução

O desenvolvimento deste sistema envolveu passos importantes como a leitura e tratamento dos dados históricos e de previsão, a implementação dos modelos de desenvolvimento da alternariose do brócolo e de *Delia radicum*, a construção das cartas de risco para estes dois inimigos e a integração das mesmas na plataforma pública de apoio aos empresários agrícolas da cultura do brócolo. Para cada um destes passos, foram construídos um ou mais subsistemas de suporte, que serão apresentados nesta secção.

3.2.1 Leitura e tratamento dos dados

Para que seja possível a construção de cartas de risco, é necessário alimentar os modelos de previsão com dados meteorológicos fiáveis. Para esta solução, foram utilizados os dados meteorológicos do Instituto Português do Mar e da Atmosfera (IPMA), que fornece informação meteorológica histórica e de previsão. Existem 137 estações meteorológicas do IPMA em Portugal, dispersas pelo território continental (figura 3.2), pelo arquipélago dos Açores (figura 3.3) e pela Região Autónoma da Madeira (figura 3.4), embora algumas destas se encontrem inoperacionais.

Foi implementado um grupo de componentes que em conjunto têm a responsabilidade de recolher e tratar os dados recolhidos. São estes:

- InfoGetter - Trata os dados recolhidos antes de serem alimentados aos modelos
- DBConnector - Efetua toda a comunicação com a base de dados do projeto
- IPMAConnector - Conecta com a API do IPMA para recolher os dados meteorológicos mais recentes

Data	Hora	T	RH	P(precipitação)	Precipitação medida
2018-09-03	15	24.6 °C	65.50%	12%	N/A
2018-09-03	16	23.7 °C	?	8.3%	N/A
2018-09-03	17	22.8+C	72.83%	4.67%	N/A

Tabela 3.1: Dados meteorológicos antes da interpolação

atual e os próximos dias. Este componente comunica com o componente *IPMAConnector* para fazer download destes dados, que estão no formato [JSON](#)

Obter o risco calculado para uma dada estação, às 23 horas do dia anterior Existem modelos que utilizam a hora anterior para calcular o risco da hora atual de um dado dia. Para que seja possível calcular o risco às 00h de cada dia com esses modelos, foi necessário armazenar uma série temporal com o risco às 23 horas do dia anterior. Esta função comunica com o componente *DBConnector* para obter este risco a partir da base de dados da plataforma.

Obter uma lista de todas as estações meteorológicas A base de dados contém uma lista com todas as estações meteorológicas. Esta função comunica com o componente *DBConnector* para obter essa lista e passá-la aos componentes que dela necessitem.

Obter uma estrutura de dados com o risco mais atual para uma dada estação Depois de lidos os ficheiros [JSON](#) com os dados meteorológicos observados no dia anterior e os previstos para os próximos dias, é feita uma organização dos mesmos num [DataFrame](#). Para cada hora de cada dia, são guardadas as informações:

- Probabilidade de precipitação - para o caso do dia ser o dia atual ou um dia futuro
- Precipitação medida - para o caso do dia ser o dia anterior
- Temperatura média¹
- Humidade relativa média¹
- VPD (*Vapour-pressure deficit*)¹

Esta estrutura de dados é então fornecida ao componente que vai executar os modelos.

Interpolarmos os dados meteorológicos que estão em falta Por vezes, os dados obtidos do IPMA têm registos vazios (por exemplo, não foi registada a probabilidade de precipitação para uma dada hora futura). Para lidar com estes casos, é feita a interpolação linear dos dados em falta, usando os dados do mesmo dia que estão presentes nos dados. As tabelas 3.1 e 3.2 mostram um exemplo antes e depois da interpolação.

¹ Este valor pode ser um valor verificado (dia anterior) ou previsto (dia atual e futuros)

Data	Hora	T	RH	P(precipitação)	Precipitação medida
2018-09-03	15	24.6 °C	65.50%	12%	N/A
2018-09-03	16	23.7 °C	69.17%	8.3%	N/A
2018-09-03	17	22.8+C	72.83%	4.67%	N/A

Tabela 3.2: Dados meteorológicos depois da interpolação

A coluna “Precipitação medida” é sempre N/A porque os dias encontravam-se no futuro quando o exemplo foi construído.

A coluna VPD não está presente pois só é calculada depois da interpolação ter sido concluída.

Depois dos dados lidos e tratados, estes são alimentados aos modelos de risco implementados.

3.2.1.2 Componente DBConnector

Este módulo é o responsável por toda a comunicação com a base de dados e foi desenvolvido com o intuito de simplificar o acesso à mesma.

Das diversas funcionalidades que implementa, destacam-se as seguintes:

Obter a lista de todas as estações Este componente comunica com a base de dados do projeto para recolher informação sobre cada uma das estações registadas

Atualizar o valor do risco A partir do novo valor, o id da estação, o modelo e a data, este componente realiza a atualização de um dado valor de risco na base de dados

Obter o risco diário para um dado dia e uma dada estação Dado o dia para o qual se quer saber o risco, o modelo e a estação, o componente comunica com a base de dados para apresentar o risco associado.

3.2.1.3 Componente IPMAConnector

Este componente é o responsável por comunicar com a API do IPMA para obter os dados meteorológicos. Disponibiliza as seguintes funcionalidades:

Obter os dados meteorológicos mais recentes Todos os dias, o IPMA disponibiliza os dados meteorológicos mais atualizados, quer seja os medidos (do dia anterior) como os de previsão (do dia atual e seguintes). Com esta funcionalidade, é possível recolher esses dados para uma pasta local.

Obter os dados medidos mais recentes Utilizada pela funcionalidade acima descrita, esta recolhe apenas os dados já medidos para uma pasta local.

Obter os dados de previsão mais recentes Também utilizada pela funcionalidade mais geral “Obter os dados meteorológicos mais recentes”, esta recolhe os dados de previsão disponibilizados pelo IPMA mais recentes.

3.2.2 Implementação e execução dos modelos de risco

Os modelos de risco implementados nesta solução foram:

HDC FV53b O modelo desenvolvido por R. Kennedy e A.M. Graham permite determinar se existe infecção ou esporulação com base na duração do período de folha molhada ou no número de horas em que existiu uma alta humidade

Modelo genérico de infecção para patologias fúngicas foliares com soma de horas molhadas

Modelo baseado no modelo genérico de infecção para patologias fúngicas foliares, mas que soma o número de horas molhadas e compara com um limiar necessário para que ocorra a infecção. Este modelo é apresentado na figura 2.4

Modelo genérico de infecção para patologias fúngicas foliares com percentagem de horas molhadas

Modelo baseado no modelo genérico de infecção para patologias fúngicas foliares, mas que determina um limiar de infecção e calcula quanto é que cada hora molhada contribui para esse limiar. Quando o limiar é atingido, a infecção ocorre. Este modelo é apresentado na figura 2.5

Modelo Delia radicum Modelo que utiliza o modelo SafeBrócolo I para calcular os graus-dias acumulados até um certo dia e com base nesse valor indica a fase atual do ciclo de vida da mosca da couve

Estes modelos foram apresentados nos capítulos 2.2 e 2.3.2.

Para que seja possível a execução dos modelos foram construídos dois componentes: RiskUpdater e DBCorrector.

3.2.2.1 Componentes RiskUpdater e DBCorrector

O componente RiskUpdater é um dos mais importantes do sistema, já que é este que coordena a execução de cada um dos modelos para cada uma das estações. Destacam-se as funcionalidades:

Adicionar um modelo Este componente guarda uma lista de modelos a executar e é possível adicionar novos modelos à lista

Executar os modelos existentes Percorre a lista de modelos (que é configurável) e executa cada um N vezes, sendo N um valor pré-determinado que varia consoante o modelo. Para isso utiliza a funcionalidade Executar um modelo N vezes

Executar um modelo N vezes O período de folha molhada é uma variável importante para os modelos de infecção por Alternaria. Uma hora é considerada molhada se, utilizando o método de folha molhada associado ao modelo, o resultado for positivo ou se a coluna correspondente à precipitação medida seja maior que 0. Isto trouxe um problema para os dados que correspondiam a datas futuras e portanto

não tinham ainda precipitação medida; em vez disso, estes dados têm um valor de probabilidade de precipitação associado. Nestes casos, foi gerado um número aleatório $0 \leq n \leq 100$ e comparado com a probabilidade de precipitação.

$$\begin{cases} \text{hora seca} & n > P(\text{precipitação}) \\ \text{hora molhada} & n \leq P(\text{precipitação}) \end{cases}$$

Para os modelos que permitem a previsão de valores de risco, é necessário que seja feita uma simulação Monte Carlo, ou seja, é necessário executar os modelos um número grande de vezes. Neste sistema, todos os modelos de previsão são executados 100 vezes para cada estação.

Atualizar o risco de um dado dia na base de dados Depois de serem executados os modelos N vezes para cada estação, é obtido o risco máximo de cada dia e esse valor passa a ser o risco previsto para o dia, relativo a uma dada estação e um dado modelo. Esta funcionalidade comunica com o componente DBConnector e atualiza na base de dados esse risco.

O componente DBCorrector tem as mesmas funcionalidades que o RiskUpdater mas serve para reescrever todos os valores de risco. O componente lê todos os ficheiros históricos com dados meteorológicos e recalcula todos os índices de risco, substituindo os valores presentes na base de dados.

3.2.2.2 Extensibilidade do sistema de execução de modelos

Um fator tido em conta aquando da implementação deste subsistema foi a extensibilidade do mesmo e, portanto, a solução final permite que se adicionem e executem novos modelos sem ter que alterar o sistema original. Para isso, basta criar uma nova classe do tipo `Model` que implemente as seguintes funções:

init Nesta função é criado o novo modelo, indicando os seguintes parâmetros:

- Nome do modelo - O nome do modelo
- Identificador do modelo - Identificador único, usado na base de dados da plataforma para identificar o modelo
- É binário ou não - Um modelo é binário quando o resultado da sua execução resulta num conjunto cujos elementos só tomam o valor 0 ou 1.
- É um modelo horário ou não - Um modelo horário é um modelo que produz risco para cada hora
- Função de agregação - Usada nos modelos horários, a função de agregação serve para calcular o risco de um dado dia com base no risco das horas do dia

longitude	latitude	estação	risco
-9,5	42.2	610	0.4
-9.49	42.2	610	0.4
-9.48	42.2	610	0.4

Tabela 3.3: Exemplo dos dados presentes no mapa de valores de risco

execute A função `execute` recebe um conjunto de dados, o último risco conhecido (às 23 horas do dia anterior) e o id da estação para a qual se quer calcular o risco. A partir destas informações executa o modelo, retornando os mesmos dados mas com uma coluna extra, a do risco.

Depois de executado cada modelo, são inseridos os valores de risco na base de dados do projeto, identificados com a data e o modelo associados ao valor, entre outras informações (listadas na secção O).

3.2.3 Construção das cartas de risco

O subsistema responsável por construir as cartas de risco funciona utilizando os valores de risco calculados no passo anterior, depois da execução dos modelos de risco para todas as estações do IPMA existentes em Portugal.

É criado um mapa (tabela 3.3) em que constam as coordenadas do território português e os valores de risco para cada conjunto de coordenadas (latitude, longitude). O risco de um conjunto de coordenadas é igual ao risco da estação meteorológica mais próxima. As coordenadas têm uma diferença de 0.01° entre elas, o que resulta numa matriz de 340×530 pontos geográficos do território português.

A partir desta matriz é gerado um raster. Este raster é então cortado para ter a forma de Portugal Continental e colorido de acordo com a escala relativa ao risco de infeção por *Alternaria* (tabela 3.4) ou ao risco de ataque por mosca da couve (tabela), conforme o modelo ao qual os dados estão relacionados.

Intervalos de risco	Cor
0.0 a 0.50	Verde
0.50 a 1.0	Verde amarelado
1.0 a 1.5	Amarelo
1.5 a 2.0	Laranja
2.0 a $+\infty$	Vermelho

Tabela 3.4: Escala de risco de alternaria

Intervalos de risco	Cor
0.0 a 0.25	Verde
0.25 a 0.50	Verde amarelado
0.50 a 0.75	Amarelo
0.75 a 1	Laranja
1 a $+\infty$	Vermelho

Tabela 3.5: Escala de risco de ataque por mosca da couve

O resultado é um conjunto de três artefactos:

- um ficheiro no formato TIFF (dados geográficos)
- um ficheiro CSV com a matriz de coordenadas e os respetivos valores de risco
- uma imagem no formato PNG, a carta de risco.

A imagem gerada é então colocada numa pasta própria, pronta a ser integrada na plataforma pública online e interpretada pelos empresários agrícolas.

3.2.4 Integração das cartas de risco na plataforma

A plataforma online construída no âmbito do projeto SafeBrócolo permite aos empresários agrícolas um maior controlo sobre as suas [parcelas](#).

Para atingir um dos principais objetivos deste projeto, foram integradas nesta plataforma cartas de risco de alternariose e da mosca da couve. Desta forma, é possível observar o risco de ocorrência destes dois inimigos em qualquer parte do território continental português. A figura 3.5 mostra a interface que permite aos empresários agrícolas visualizar o risco de alternariose e as cartas de risco.

As imagens das cartas de risco são obtidas através de uma pasta própria onde foram colocadas pelo subsistema gerador de rasters.

O painel de controlo do utilizador na plataforma pública possibilita a visualização das seguintes cartas de risco:

- Estações meteorológicas:
 - Risco de *Alternaria* (modelo HDCFV53b)
 - Risco de ataque da mosca da couve (modelo *Delia radicum*)
- Satélite:
 - Risco *Alternaria*
 - Risco *Delia radicum* (larvas)
 - Risco *Delia radicum* (adultos)

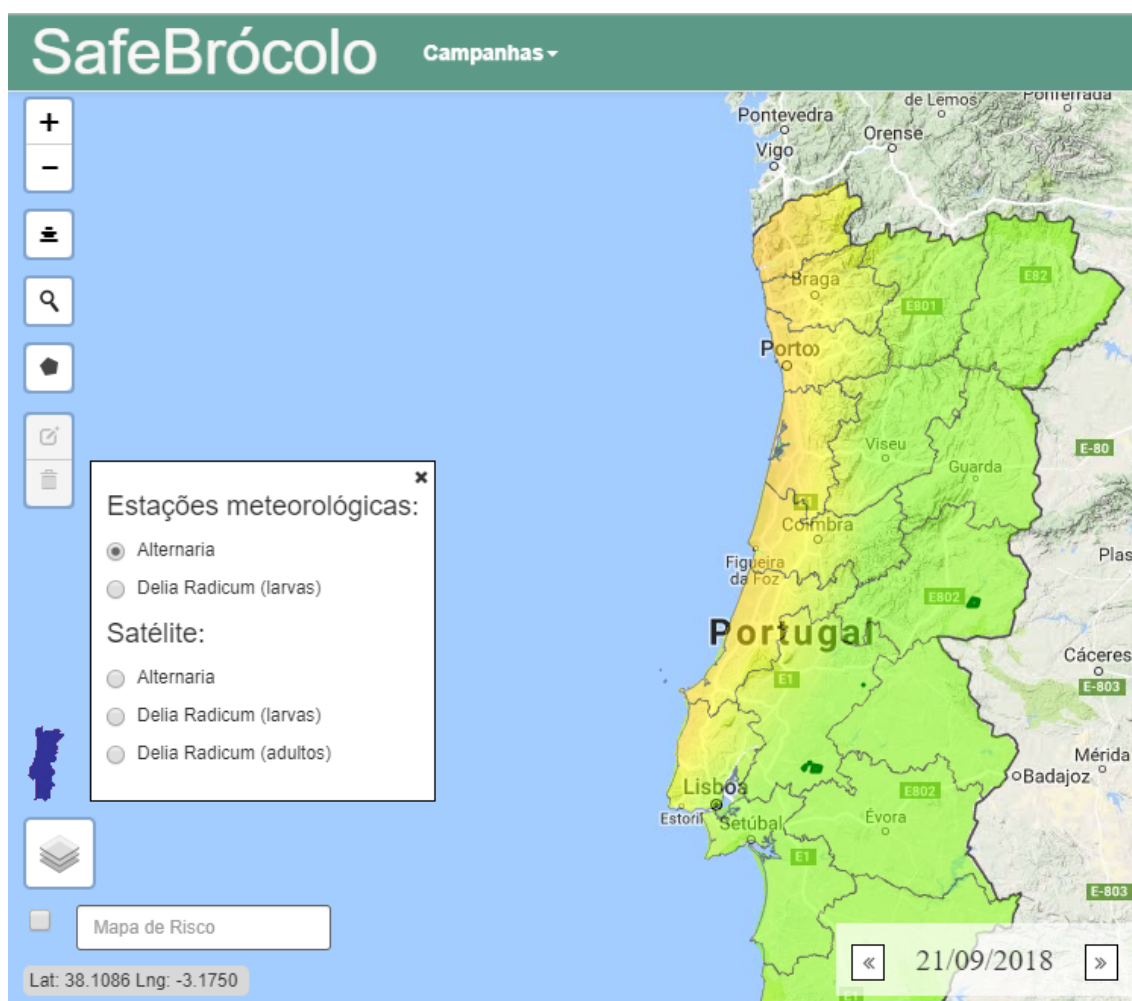


Figura 3.5: Carta de risco de alternariose apresentada pela plataforma

Os modelos de *Delia radicum* do satélite recorrem a dados de temperatura de superfície do solo (LST) obtidos a partir do satélite EUMETSAT (com limiar mínimo de 6°C e limiar máximo de 30°C) e compensam dados inexistentes por cobertura de nuvens através das estações do IPMA mais próximas.

O modelo de risco de alternaria do satélite assume que dias de reduzida amplitude térmica, com temperatura acima dos 10°C, correspondem a dias de humidade elevada e logo de condições adequadas ao desenvolvimento de fungos.

Além das cartas de risco, também foram incluídos os valores de risco para o dia anterior, o dia atual e para 4 dias no futuro. Estes valores, bem como um gráfico histórico dos valores de risco para a parcela, podem ser visualizados na página de perfil de cada parcela (figura 3.6) e permitem que os empresários agrícolas antecipem ataques dos inimigos mencionados e ajam em conformidade, potencialmente salvando hectares de cultura.

Como extra, foi desenvolvida uma página para análise histórica dos valores de risco para cada estação (figura 3.7). Esta página permite ver os valores de risco e filtrar por

Risco para Cultura de brócolos



Figura 3.6: Página com o risco associado a uma parcela

intervalo de tempo (1 semana, 2 semanas, 1 mês, 2 meses ou todos os dados). Apresenta também um gráfico com o risco para cada dia e o risco acumulado de 3 dias. Desta forma, é possível extrair padrões para uma determinada zona geográfica e prever potenciais ataques.

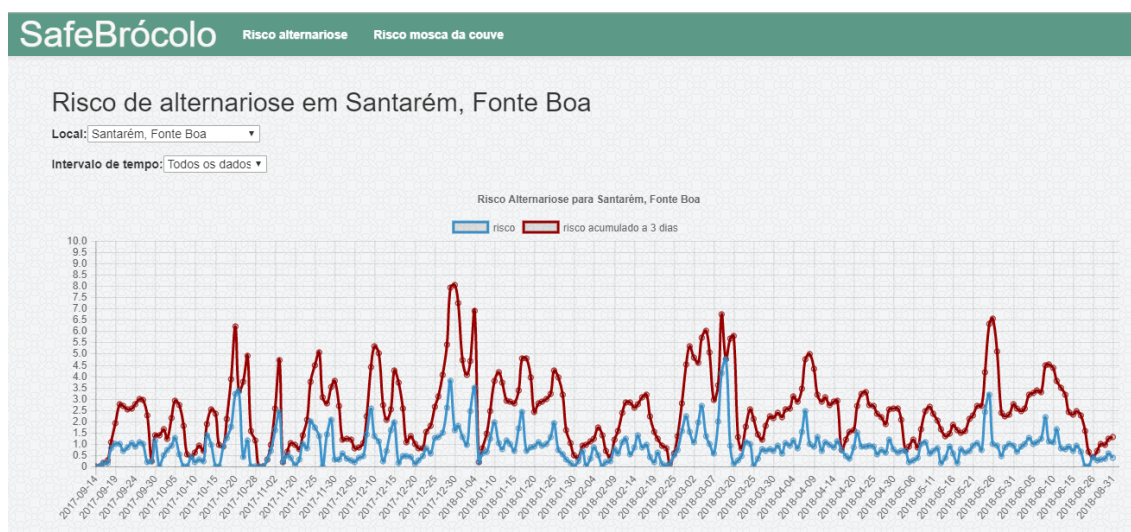


Figura 3.7: Página que permite visualizar o histórico do risco de todas as estações meteorológicas

3.3 Modelo de dados

Para dar apoio ao sistema de cálculo de risco, foram criadas relações na base de dados já existente da plataforma online. Na figura 3.8 é possível visualizar o diagrama de base de dados da solução, que inclui apenas as novas relações e a tabela campanhaemsector, que forma a ligação com o resto do esquema da base de dados. Esta tabela representa uma dada parcela numa campanha.

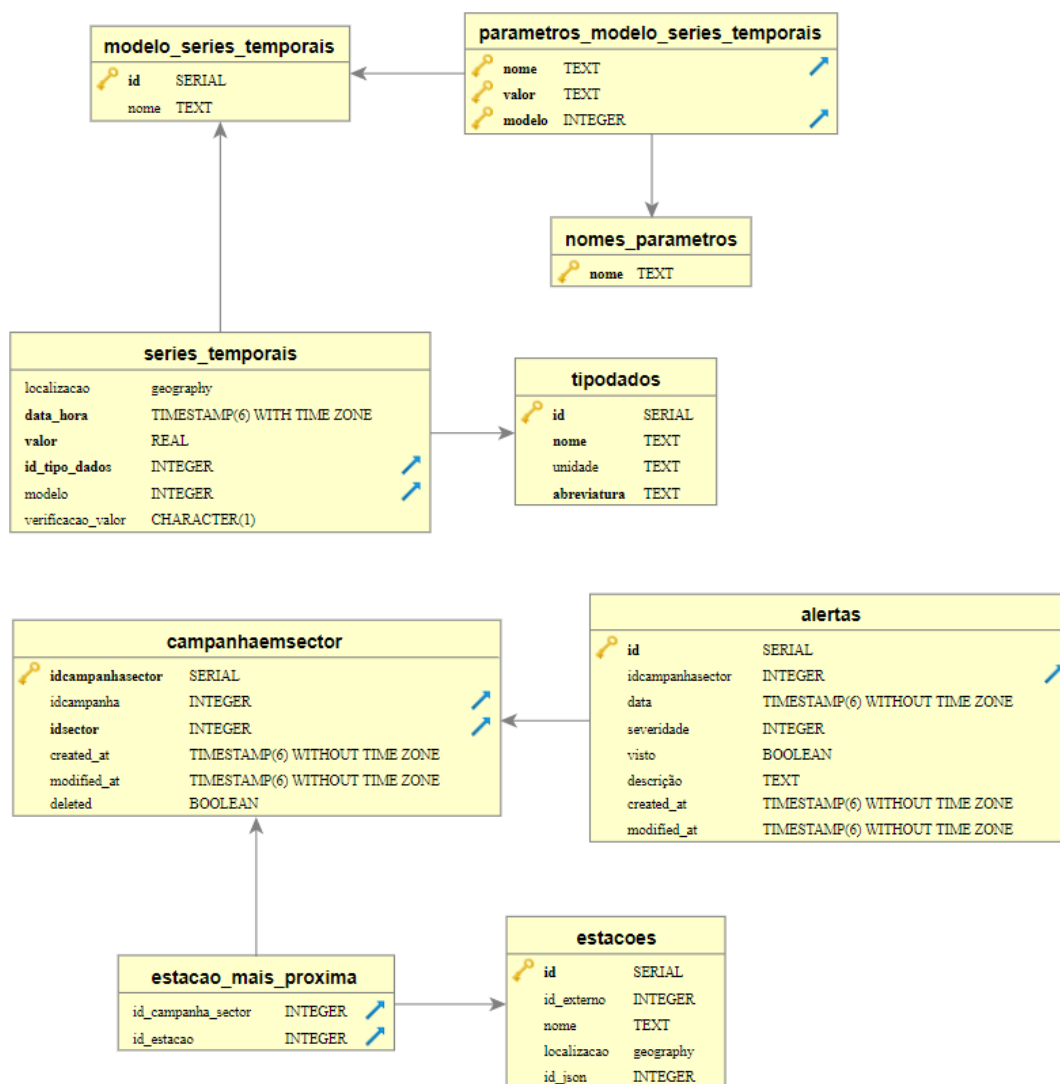


Figura 3.8: Modelo de dados da solução

3.3.1 Relação estacoes

A tabela **estacoes** serve para guardar informação acerca de cada uma das estações meteorológicas utilizadas. Contém as seguintes colunas:

id um identificador único e interno associado à estação

id_externo o identificador da estação na fonte de informação

nome o nome da estação

localizacao a localização geográfica da estação, armazenada no formato *geography*.

id_json um outro identificador global usado para identificar uma estação no serviço do IPMA. Este identificador é usado para recolher os dados de cada estação, já que os dados estão disponíveis num *endpoint* no formato:

`http://.../id_json.json`

em que `id_json` é o identificador representado por esta coluna na relação.

3.3.2 Relação `estacao_mais_proxima`

A tabela `estacao_mais_proxima` armazena a estação mais próxima de cada uma das parcelas registadas na plataforma. Quando uma parcela é adicionada, é automaticamente calculada a estação que se encontra mais perto e adicionado um registo nesta relação. As colunas presentes nesta tabela são:

`id_campanha_sector` um identificador único e interno associado à parcela

`id_estacao` o identificador da estação mais próxima. Corresponde à coluna `id_externo` da relação `estacoes`

3.3.3 Relação `series_temporais`

A tabela `series_temporais` têm o propósito de armazenar séries de dados que tenham potencial para variar ao longo do tempo. É nesta relação que são armazenados os valores do risco de alternariose e mosca da couve para cada um dos dias e cada uma das estações meteorológicas. Contém as seguintes colunas:

`localizacao` a localização associada ao valor

`data_hora` o momento associado ao valor

`valor` o resultado da medição/cálculo

`id_tipo_dados` o tipo de dados a que corresponde o valor

`modelo` o modelo usado para calcular o valor

`verificacao_valor` 'P' se o valor corresponder a uma previsão ("Previsto") e 'V' se o valor tiver sido calculado com base em medições reais ("Verificado")

3.3.4 Relação `modeloserie_temporais`

Esta relação contém os modelos usados para calcular os valores armazenados na tabela `series_temporais`. É composta pelas colunas:

`id` o identificador do modelo

`nome` o nome do modelo

3.3.5 Relações `parametros_series_temporais` e `nomes_parametros`

Estas duas relações servem de suporte à relação `modelo_series_temporais`, armazenando os parâmetros usados em cada modelo. A relação `parametros_series_temporais` tem as colunas:

nome o nome do parâmetro

valor o valor usado no parâmetro

modelo o modelo associado a este parâmetro

A relação `nomes_parametros` contém apenas uma coluna com o nome do parâmetro.

3.3.6 Relação `tipodados`

A relação `tipodados` contém a definição das diferentes séries de dados armazenadas na relação `series_temporais`. As três séries atualmente existentes são:

- O risco acumulado (3 dias) - Tanto o risco de alternariose como de infestação por mosca da couve acumulados têm este tipo de dados associado
- O risco não-acumulado (1 dia) - O risco diário de alternariose ou mosca da couve
- O risco às 23 horas de um dado dia

Esta relação contém os atributos:

id um identificador único e interno associado a um tipo de dados

nome uma pequena descrição do tipo de dados

unidade as unidades da série (por exemplo, numa série de temperaturas, a unidade poderia ser graus Celsius)

abreviatura a abreviatura da série

3.3.7 Relação `alertas`

A tabela `alertas` contém os alertas gerados pelo sistema, que avisam os empresários agrícolas do potencial risco iminente de ataque por parte de *Alternaria* ou *Delia radicum*. A plataforma online lê os registos inseridos nesta estrutura de dados e apresenta ao utilizador os que ainda não tenham sido vistos. As colunas presentes nesta tabela são:

id um identificador único e interno associado ao alerta

idcampanhasector o identificador da parcela à qual o alerta está associado

data a data à qual o alerta corresponde

severidade a severidade do alerta, neste caso, a categoria do risco (risco "baixo", "médio" ou "alto")

visto esta coluna fornece informação sobre a visualização de um alerta. Toma o valor `true` quando o alerta já foi visto pelo utilizador da plataforma online e `false`, caso contrário.

descrição O texto apresentado ao utilizador quando o alerta é visualizado

created_at um identificador único e interno associado à parcela

modified_at um identificador único e interno associado à parcela

3.4 Tecnologias utilizadas

A solução desenvolvida está maioritariamente desenvolvida na linguagem de programação Python. Para que o desenvolvimento da solução fosse possível, foram utilizadas várias bibliotecas, das quais se destacam três:

SciPy A biblioteca SciPy [24] disponibiliza várias funcionalidades associadas à matemática, estatística, ciência e engenharia. Foi utilizada para criar funções de interpolação e aplicá-las aos dados meteorológicos, de modo a preencher as lacunas existentes nos mesmos

Pandas A biblioteca Python mais utilizada foi sem dúvida Pandas, que fornece ferramentas de análise de dados e permite criar estruturas de dados complexas mas que mantém uma alta performance. As estruturas de dados que fornece (designadas *dataframes*) foram utilizadas para guardar os dados meteorológicos e passá-los entre componentes da solução

GDAL Esta biblioteca tem o propósito de gerar rasters a partir de dados geograficamente identificáveis. Foi utilizada para a geração das cartas de risco a partir dos valores de risco para cada coordenada do território português.

A existência destas bibliotecas facilitou o desenvolvimento dos vários componentes da solução.

3.5 Desafios encontrados e soluções

A implementação da solução apresentada incluiu alguns desafios, tanto na leitura e tratamento dos dados meteorológicos recolhidos, como na implementação dos próprios modelos e por fim na construção e integração das cartas de risco. Nesta secção apresentam-se as soluções encontradas para esses desafios.

3.5.1 Leitura e tratamentos dos dados

Por vezes, as estações meteorológicas que recolhem os dados por todo o país tem uma falha momentânea ou qualquer outro problema que impeça a recolha dos dados na totalidade. Por exemplo, os dados meteorológicos podem ter certas horas sem o valor da temperatura média do ar por motivos de avaria.

Para combater este problema, é usada uma interpolação polinomial e os dados são tratados antes de serem alimentados aos modelos, de forma a preencher as lacunas que possam existir.

São criadas 3 funções de interpolação empírica polinomiais, com base nos dados que estão completos, dentro dos dados recebidos para uma dada estação e dia. As funções são:

- Função de interpolação polinomial para a temperatura média do ar
- Função de interpolação polinomial para a humidade relativa
- Função de interpolação polinomial para a precipitação

Estas funções são geradas automaticamente pela biblioteca Scipy, sendo que são diferentes para cada conjunto de dados. São depois utilizadas para obter os valores em falta, tendo em conta os valores que estão presentes, completando assim o conjunto de dados.

3.5.2 Implementação dos modelos de risco

A maioria dos modelos necessita de obter informação extra sobre uma dada hora, como por exemplo, se durante uma dada hora, a folha esteve molhada. Como as estações meteorológicas do IPMA não têm sensores de folha molhada, cada um dos modelos é obrigado a utilizar um dos métodos apresentados no capítulo 2.4 para determinar essa informação.

Para facilitar o acesso a esta informação extra e uniformizar o cálculo da mesma foi criado o módulo Python HourData que representa uma dada hora num conjunto de dados meteorológicos. Entre outras funcionalidades, este módulo disponibiliza a função `is_leaf_wet` que informa se numa dada hora a folha esteve molhada, permitindo especificar que método de folha molhada se pretende aplicar.

3.5.3 Construção das cartas de risco

A construção de cartas de risco implica o uso de bibliotecas para construção e recorte de rasters. Como mencionado no capítulo 3.4, para esta solução foi usada a biblioteca GDAL. Mais especificamente foram usados quatro comandos diferentes: `gdal_grid`, `gdalwarp`, `gdaldem` e `gdal_translate`.

O comando `gdal_grid` cria uma grelha de dados com base numa fonte de informação, neste caso, um ficheiro csv com o risco para cada coordenada. É necessário fornecer, entre outros parâmetros, as dimensões pretendidas da grelha resultante (340x530).

```
1 gdal_grid -a invdist:smoothing=0.1 -a_srs EPSG:4326 -txe -9.5 -6.1 -tye 36.9  
42.2 -outsize 340 530 -of GTiff -ot Float32 -l risco risco.vrt risco2.tiff
```

O comando `gdalwarp` recorta a grelha de acordo com um ficheiro shapefile (de forma) com a forma do território português.

```
1 gdalwarp -dstnodata -999 -dstalpha -cutline pt10m.shp risco2.tiff risco.tiff
```

O comando `gdaldem` aplica uma escala de cores à grelha, de acordo com os valores de cada célula e um ficheiro de escala. Neste caso, foram usados dois ficheiros de escala diferentes: um para a escala de risco para a alternariedade e outro para a mosca da couve.

```
1 gdaldem color-relief -alpha risco.tiff colors_filename nome_final_ficheiro_tiff
```

Finalmente, o comando `gdal_translate` cria uma imagem com o formato `.png` a partir do raster gerado.

```
1 gdal_translate -of PNG -scale nome_final_ficheiro_tiff nome_final_ficheiro_png
```

3.5.4 Conclusão

Foi construída uma solução que lê dados meteorológicos, trata-os e executa modelos de risco. Utiliza então os resultados dos modelos para gerar cartas de risco que são disponibilizadas para a plataforma online.

Foi necessário estender a base de dados existente para incluir os dados relativos às estações meteorológicas e aos resultados dos modelos de risco, bem como os alertas a mostrar aos empresários agrícolas.

A existência de ferramentas como o SciPy, Pandas e GDAL facilitou o desenvolvimento desta solução.

A criação de uma solução que utiliza dados de uma fonte externa implica o tratamento de lacunas que possam ocorrer nos mesmos. Este tratamento foi feito com o uso de funções de interpolação linear.

Além disso, foi necessário colmatar a falta de sensores de folha molhada nas estações meteorológicas do IPMA. Para isso foram estudados e utilizados modelos de folha molhada.

RESULTADOS OBTIDOS

Neste capítulo serão apresentados os resultados do desenvolvimento do projeto, sendo incluída uma análise dos modelos e um estudo do desempenho da plataforma, quer em termos de tempo como de espaço.

4.0.1 Modelos de folha molhada

As estações meteorológicas do IPMA não têm sensores de folha molhada, portanto não é possível obter medições exatas da duração do período de folha molhada no local. Esta é uma variável importante para a determinação da ocorrência de infeção pelo que a falta desta informação é considerada um problema.

Para o resolver, foram analisados e comparados os modelos de folha molhada mencionados na literatura e apresentados na secção 2.4. Foi utilizada uma estação meteorológica com sensor de folha molhada, localizada na Chamusca, e comparados os dados históricos recolhidos nessa estação com a aplicação de cada um dos métodos de folha molhada.

Foram testados os modelos:

- CART/SLD
- Limiar de VPD constante
- Limiar de humidade relativa constante
- Limiar de DPD constante

A opção escolhida foi o modelo que usa o VPD pois além da boa adesão aos dados do sensor, já foi usado como indicador de folha molhada na literatura referida, no desenvolvimento de modelos de infeção e esporulação de *A. brassicae* e *A. brassicicola* em culturas de brócolo.

		Sensor		Total
		Folha molhada	Folha seca	
VPD	Folha molhada	4273 (34.4%)	2128 (17.1%)	6401
	Folha seca	303 (2.4%)	5720 (46.1%)	6023
Total		4576	7848	12424

Tabela 4.1: Matriz de confusão para estimativa de folha molhada a partir de défice de pressão de vapor (DPV) para a estação da Chamusca (setembro 2015 a janeiro de 2017)

Na tabela 4.1 é possível observar a matriz de confusão com os resultados da comparação entre utilização do VPD para determinar se uma dada hora é molhada e a utilização de um sensor local.

Como se pode observar, o modelo que usa o VPD consegue determinar em 34,4% dos casos que uma hora é molhada quando existe realmente essa indicação por parte do sensor local. Em 17,1% dos casos, o modelo diz que a folha está molhada quando na realidade não está.

Isto pode depender de fatores como a calibração do sensor mas no geral não é um problema porque a infeção necessita de folha molhada para ocorrer. Mais importante, o modelo só diz que a folha está seca quando na realidade esta está molhada em 2.4% dos casos. No geral, o modelo do VPD coincidiu com a realidade em 80.5% dos casos.

4.1 Validação dos modelos

Durante as campanhas de outono de 2015, primavera de 2016 e de outono de 2016, a equipa do projeto Safebrócolo colheu amostras de brócolos com sintomas de alternariose de uma cultura de brócolos localizada na zona da Chamusca (Santarém). Nesta zona foi também instalada uma estação meteorológica para recolher dados climáticos.

Os dados foram introduzidos nos modelos, que depois de executados, forneceram níveis de risco para cada dia. Nos modelos mais corretos, as datas em que foram recolhidas amostras infetadas com *Alternaria* deverão estar próximas das datas para as quais os modelos apontam um risco significativo, tendo sempre em conta que o aparecimento dos sintomas ocorre 4 a 6 dias depois da infeção.

Os valores de risco fornecidos pelos modelos tomam valores entre 0 e $+\infty$. Considera-se risco significativo um risco maior que 1, sendo que este valor significa que as condições necessárias à infeção estão presentes.

4.1.1 Modelo 5 horas de folha molhada - M5HM

Tal como descrito no ciclo de vida dos fungos do género *Alternaria* (capítulo 1.5), a infeção ocorre quando as temperaturas estão entre os 15 e os 25°C e a folha está molhada durante cerca de 5 horas. Com base nesta informação foi criado um modelo que, quando deteta que estas condições se verificam, fornece um valor de risco igual ou superior a 1.

O algoritmo do modelo (listagem 4.1) é o seguinte: O risco começa com o valor 0. Para cada hora registada nos dados meteorológicos recebidos, se a folha esteve molhada durante essa hora e a temperatura média da hora estiver entre os 15 e os 25°C, então o risco aumenta em 0.2. Se numa dada hora alguma das condições não se verificar, então o risco volta a ser 0.

Listagem 4.1: Algoritmo do modelo M5HM

```

1 risco = 0;
2 risco_de_cada_hora = []
3 for _, hora in dados.iterrows():
4     if hora['hora_molhada'] and 15 <= hora["air_temp_avg"] <= 25:
5         risco += 0.2
6     else:
7         risco = 0
8     risco_de_cada_hora.append(risco)
9 dados['risco'] = risco_de_cada_hora

```

Este modelo base não considera a evolução da infeção para outros intervalos de temperatura nem evolução da doença dependente da temperatura.

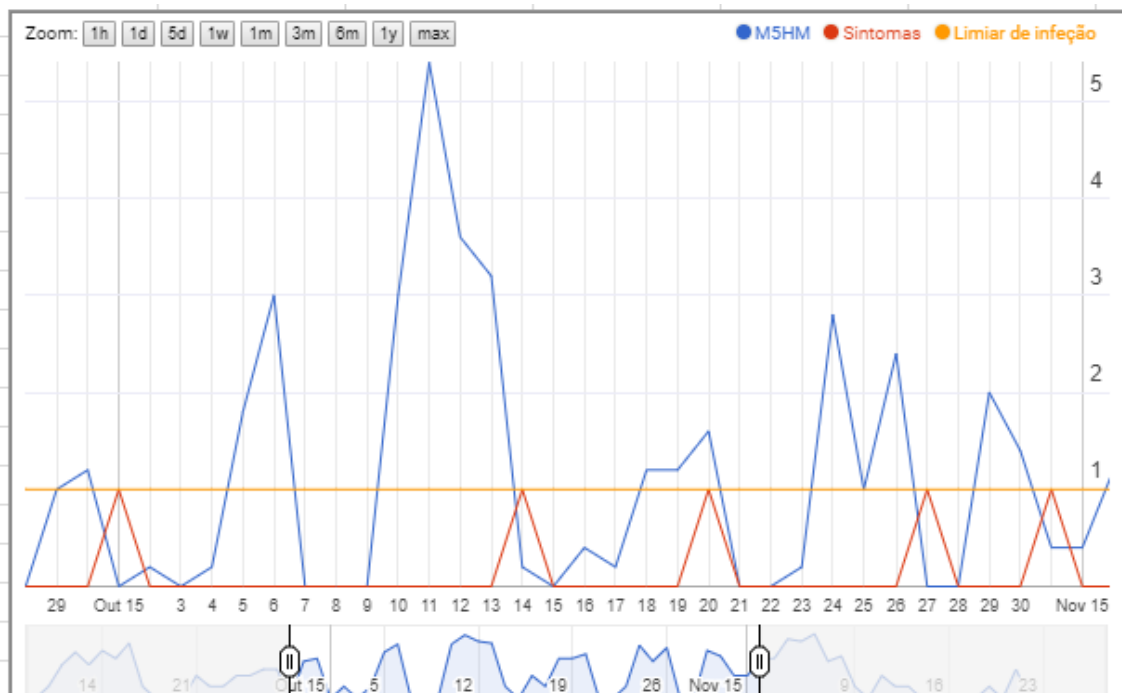


Figura 4.1: Risco de alternariose segundo o modelo M5HM para Outubro de 2015

Na figura 4.1 pode ser visualizado o risco para o mês de Outubro de 2015. A linha laranja representa o limiar de infeção (1) e os picos a vermelho indicam os dias em que foram observados sintomas na cultura.

O modelo parece ter uma boa capacidade preditiva em relação à ocorrência de infeção, mostrando valores de risco altos para os dias antes do aparecimento de sintomas.

4.1.2 Modelo HDC FV53b

O modelo introduzido no capítulo 2.2.2 foi também validado com os dados recolhidos pela equipa do projeto na cultura de brócolos na Chamusca. Recorde-se que os autores R. Kennedy e A.M. Graham não apresentam uma função para o gráfico relativo às condições necessárias para infeção incluídos no documento, pelo que foi necessário efetuar uma interpolação empírica através dos pontos conhecidos.

Foram experimentados polinómios de 2º grau, 3º grau e 4º grau para descobrir a função que melhor se adequa aos pontos. As funções encontradas encontram-se na tabela 4.2. A variável t representa a temperatura média do ar e o resultado da função é o período de folha molhada mínimo para que a infeção ocorra.

Grau	Função
2º grau	$0.068142t^2 - 3.036729t + 41.623669$
3º grau	$-0.001095285t^3 + 0.122358529t^2 - 3.865969232t + 45.465817963$
4º grau	$0.0003485713t^4 - 0.0241009885t^3 + 0.6621958390t^2 - 9.1539946225t + 63.5472257137$

Tabela 4.2: Funções encontradas por interpolação empírica dos dados conhecidos para o modelo HDC FV53b

O algoritmo do modelo (listagem 4.2) funciona da seguinte forma: O risco começa com o valor 0. Sabendo que cada hora molhada pode ter um contributo diferente para o ritmo do crescimento do risco (dependendo da temperatura média do ar nessa hora), é usada a função inversa do polinómio para calcular esse contributo e somá-lo ao risco atual. Se a hora for seca, é verificada a temperatura média do ar (T_{ar}) e o VPD durante essa hora; se $T_{ar} > 8^\circ\text{C}$ ou $VPD > 2$, então o ritmo de infeção volta a 0.

Listagem 4.2: Algoritmo do modelo HDC FV53b

```

1 risco = 0;
2 risco_de_cada_hora = []
3 for _, hora in dados.iterrows():
4     if is_leaf_wet(dados, hora, "vpd", 3):
5         risco += ritmo_crescimento_infecao()
6     else if not (hora.air_temp_avg <= 8.0 and hora.vpd <= 2):
7         risco = 0
8     risco_de_cada_hora.append(risco)
9 dados['risco'] = risco_de_cada_hora

```

Note-se que, tal como concluído no capítulo 4.0.1, foi usado o modelo do limiar do VPD com valor limite de 3 kPa para determinar se numa dada hora, a folha esteve molhada.

A figura 4.2 mostra que o risco apresentado por cada um dos polinómios é na realidade muito semelhante. A linha verde representa o limiar de infeção (1) e os picos a roxo indicam a presença de sintomas.

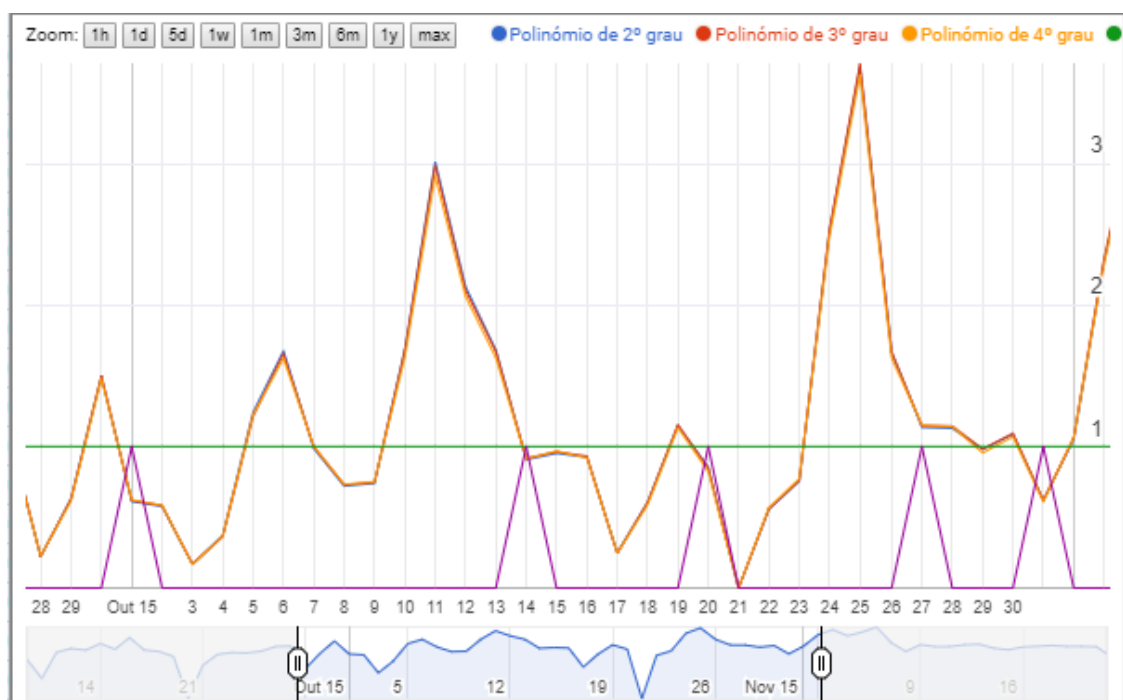


Figura 4.2: Risco de alternariose segundo o modelo HDC FV53b, comparação entre polinômios

Para evitar falsos negativos, foi usada a função que maximiza os 3 polinômios. Na figura 4.3 pode ser visualizado o risco para o mês de Outubro de 2015 utilizando esta nova função.

Este modelo aparenta conseguir detetar e prever a ocorrência de alternariose com sucesso, como se pode observar pelos valores de risco altos que ocorrem alguns dias antes dos sintomas aparecerem.

4.1.3 Modelo genérico de infecção para patologias fúngicas foliares - soma - MGIPFF-S

Desenvolvido como alternativa a uma das implementações do modelo genérico de infecção para patologias fúngicas foliares, o modelo MGIPFF-S tem o seu funcionamento descrito no diagrama da figura 2.4.

Os parâmetros do modelo foram baseados no modelo original e nos valores presentes na literatura para *Alternaria brassicae* [18, p. 4], visto não ter sido possível obter a mesma informação para *Alternaria brassicicola*. Foram utilizados os valores:

- Temperatura mínima: 2.6 °C
- Temperatura máxima: 35 °C
- Temperatura ótima: 18 °C
- Período de folha molhada mínimo: 6 horas

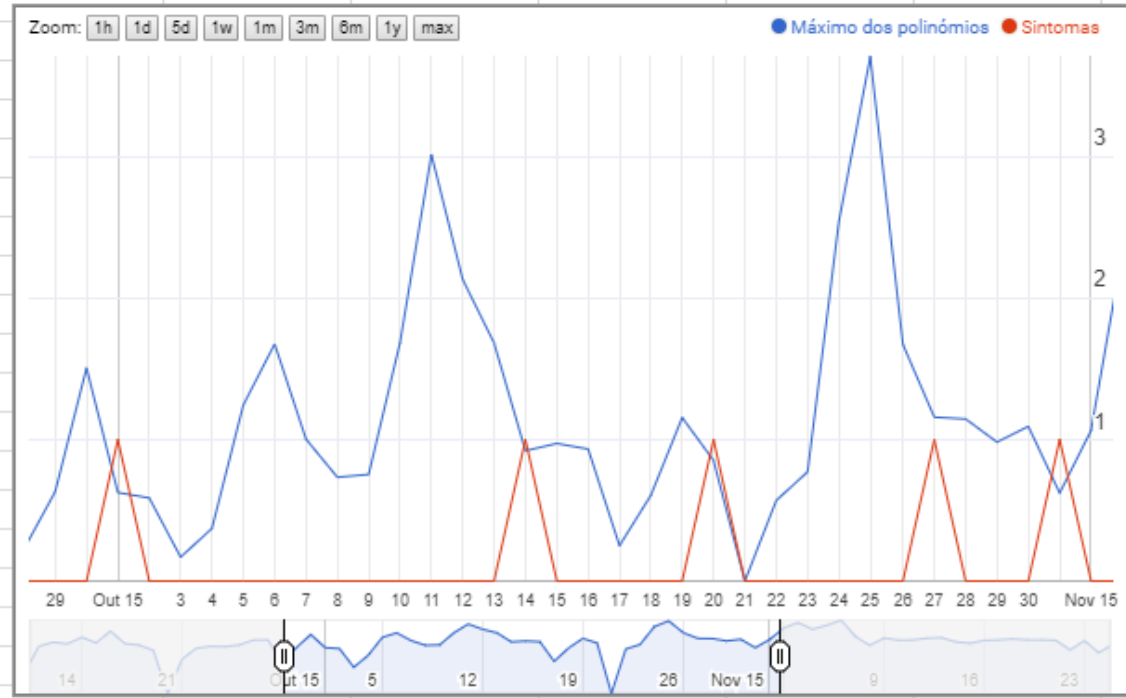


Figura 4.3: Risco de alternariose segundo o modelo HDC FV53b com o máximo dos polinômios para Outubro de 2015

- Período de folha molhada máximo: 22 horas
- Período seco máximo (d_{50}): 2 horas
- Método de folha molhada: VPD
- Valor VPD máximo para ser considerada hora molhada: 3
- Função de desenvolvimento:

$$f(T) = \left(\frac{T_{max} - T}{T_{max} - T_{opt}} \right) \left(\frac{T - T_{min}}{T_{opt} - T_{min}} \right)^{(T_{opt} - T_{min}) / (T_{max} - T_{opt})}$$

(Yin et. al [26])

Na figura 4.4 pode ser visualizado o risco para o mês de Outubro de 2015. A linha laranja representa o limiar de infecção e os picos vermelhos indicam observações de sintomas. Este modelo produz valores de maior dimensão pois o resultado do mesmo é o número de eventos de potencial infecção lançados.

Também este modelo apresenta uma boa adesão aos dados das observações e aparenta ter uma boa capacidade preditiva da ocorrência de alternariose.

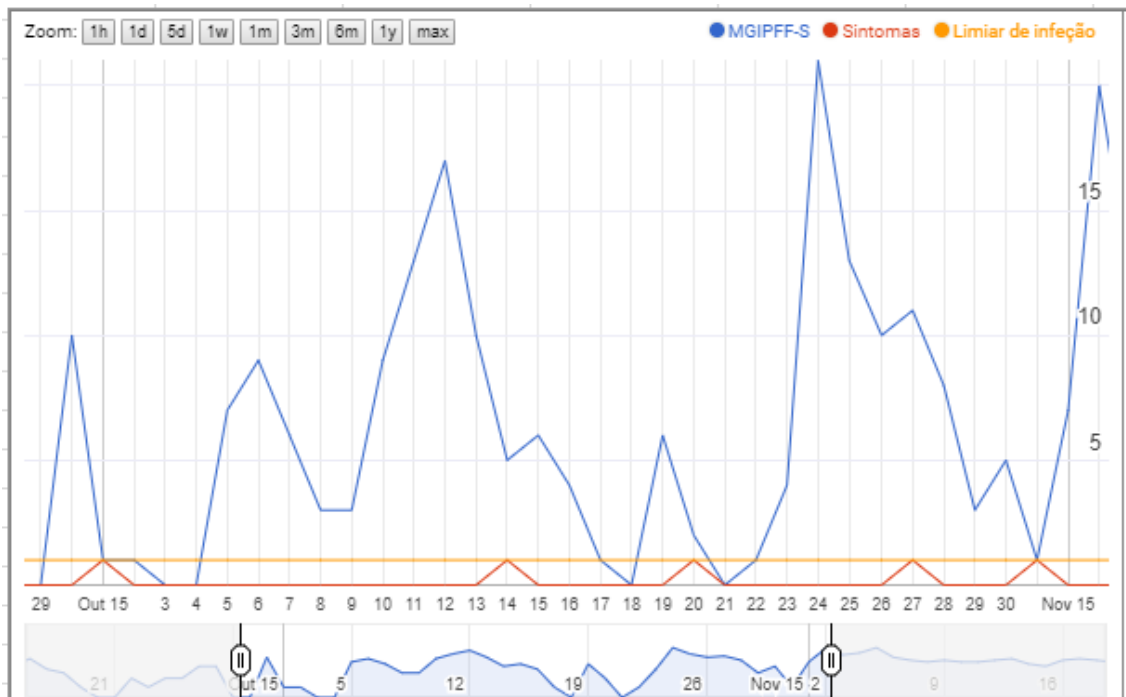


Figura 4.4: Risco de alternariose segundo o modelo MGIPFF-S

4.1.4 Modelo genérico de infecção para patologias fúngicas foliares - percentagem - MGIPFF-P

Uma outra alternativa à implementação do modelo genérico de infecção para patologias fúngicas foliares é o modelo MGIPFF-P, que, tal como descrito no capítulo 2.2, utiliza uma variável DI (Desenvolvimento da Infecção) que começa com valor 0 para todas as coortes e, em cada hora propícia (ou seja, em que $T_{min} \leq T_{ar} \leq T_{max}$ e a folha esteve molhada), aumenta em $1/W(T_h)$. $1/W(T_h)$ representa a parte que uma hora (a uma dada temperatura T_h) contribuiu para o tempo total necessário. Quando a variável atinge o valor 1, então foram reunidas as condições para a infecção ocorrer.

Os parâmetros do modelo são iguais aos do modelo MGIPFF-S:

- Temperatura mínima: 2.6 °C
- Temperatura máxima: 35 °C
- Temperatura ótima: 18 °C
- Período de folha molhada mínimo: 6 horas
- Período de folha molhada máximo: 22 horas
- Período seco máximo (d_{50}): 2 horas
- Método de folha molhada: VPD

- Valor VPD máximo para ser considerada hora molhada: 3
- Função de desenvolvimento:

$$f(T) = \left(\frac{T_{max} - T}{T_{max} - T_{opt}} \right) \left(\frac{T - T_{min}}{T_{opt} - T_{min}} \right)^{(T_{opt} - T_{min}) / (T_{max} - T_{opt})}$$

(Yin et. al [26])

Na figura 4.5 pode ser visualizado o risco para o mês de Outubro de 2015. A linha laranja representa o limiar de infecção e os picos vermelhos indicam observações de sintomas. Tal como os outros modelos, este também apresenta uma boa capacidade preditiva da ocorrência de alternariose.

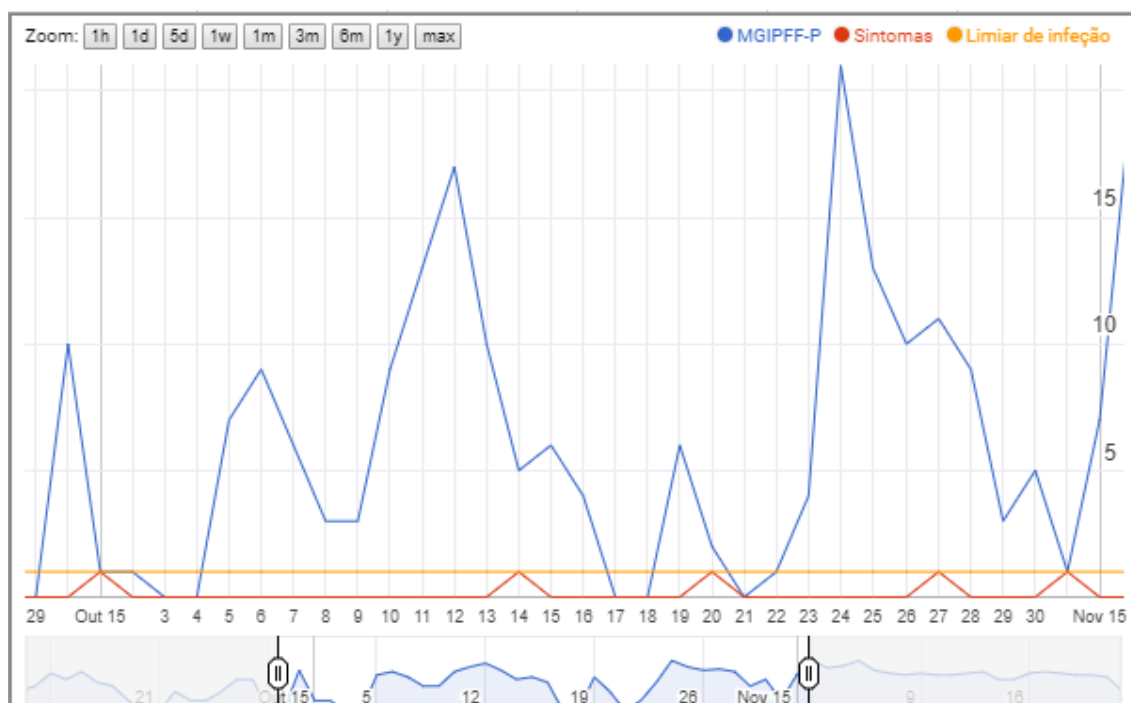


Figura 4.5: Risco de alternariose segundo o modelo MGIPFF-P para Outubro de 2015

4.1.5 Conclusões

O mês de outubro de 2015 caracterizou-se pela ocorrência de alternariose na cultura. Todos os modelos aqui apresentados conseguem indicar a ocorrência de infecção poucos dias antes de serem registados sintomas, representando uma boa adesão aos dados de registo de sintomas.

Fica apenas por explicar o alto risco calculado pelos modelos nos primeiros dias de outubro, que não coincidiu com o registo de sintomas na parcela estudada.

O projeto SafeBrócolo decidiu adotar o modelo HDC FV53b com a função do máximo dos polinómios como fonte de informação para a geração das cartas de risco que estão disponíveis na plataforma online.

4.2 Modelo Delia Radicum

O modelo de desenvolvimento da mosca da couve construído pela equipa do projeto SafeBrócolo (capítulo 2.3.2) foi usado para construir um modelo de risco para este inimigo.

Recorde-se que os insetos necessitam de um certo número de graus-dia acumulados para atingir as várias fases de desenvolvimento e que esta espécie causa mais estragos quando está na fase de larva. Este modelo começa com criar 10 instâncias de mosca da couve.

As instâncias começam por ter todas o mesmo número de graus-dia acumulados, que corresponde ao número de graus-dia acumulados até ao dia atual (desde 1 de Janeiro).

Assume-se que no dia 1 de janeiro de cada ano a população é constituída apenas por pupas, em fases de desenvolvimento uniformemente distribuídas, ou seja, a cada instância foram adicionados mais 20 graus-dia acumulados que a anterior, o que resulta numa soma de 0 até 200 graus-dia aos iniciais.

De seguida, é verificado o número de graus-dia acumulados das 10 instâncias. Por cada instância que se encontre na fase larva o risco aumenta 10%.

A figura 4.6 mostra o funcionamento das simulações de desenvolvimento e a sua comparação com os registos de sintomas de ataque de mosca da couve. Note-se que a maioria das observações de sintomas ocorre quando o modelo aponta para uma maior concentração de larvas.

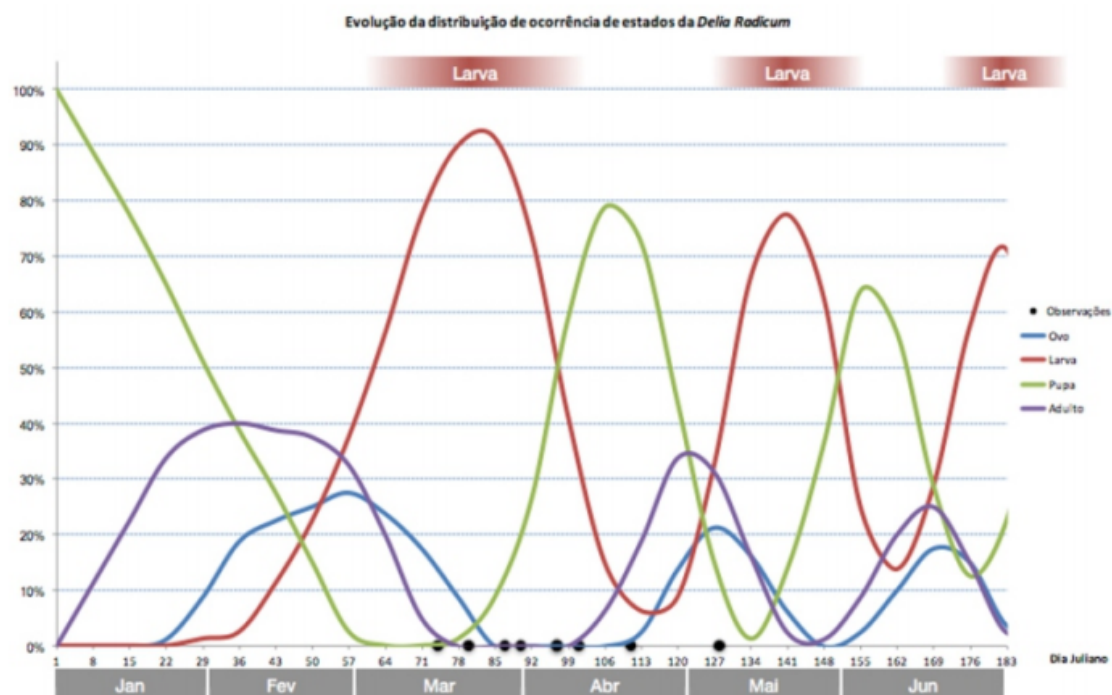


Figura 4.6: Modelo de desenvolvimento de Delia radicum - Projeto SafeBrócolo

4.3 Desempenho da solução

O cálculo do risco diário de *Delia Radicum* e *Alternaria* demora cerca de 31 minutos (média de 10 amostras) numa máquina com as seguintes especificações:

- Processador: Intel(R) Core(TM) i5 2.40GHz
- Memória: 8.00GB
- Sistema operativo: Linux Mint 18.3 (Sylvia)

Este tempo inclui o cálculo do risco para o dia anterior, o dia atual e até 4 dias depois do dia atual e a geração dos rasters para estes dias.

O espaço ocupado pela solução em disco é de 34.5 MB, excluindo os rasters. Cada raster ocupa em média 817,30 KB em disco e consiste em três artefactos diferentes: uma imagem (.png) com a carta de risco, um ficheiro no formato TIFF e um ficheiro (.csv) com os valores do risco para cada coordenada do território português continental.

4.4 Conclusões

Foram desenvolvidos e avaliados os modelos de folha molhada CART/SLD, Limiar de VPD constante, Limiar de humidade relativa constante e Limiar de DPD constante, sendo que o modelo escolhido foi o modelo de Limiar de VPD constante.

Em relação aos modelos de risco para *Alternaria* e *Delia radicum*, foram desenvolvidos e avaliados os modelos M5HM, HDC FV53b, MGIPFF-S e MGIPFF-P, sendo que o projeto SafeBrócolo adotou o modelo HDC FV53b para calcular o risco de *Alternaria* a apresentar na plataforma online e o modelo do próprio projeto para o risco de *Delia radicum*.

Os modelos desenvolvidos têm uma boa adesão com a realidade, sendo também eficientes em termos de espaço e de tempo.

CONCLUSÕES E TRABALHO FUTURO

Neste capítulo são apresentadas as conclusões tiradas ao longo do projeto bem como algumas ideias de possíveis melhorias a realizar no futuro.

5.1 Conclusões

Foi efetuado um estudo do estado da arte dos atuais modelos de desenvolvimento e previsão de risco para *Alternaria* e *Delia radicum*.

Os modelos foram avaliados e comparados com base na sua capacidade de assinalar com alto risco os dias em que realmente houveram ataques/infeções (dados históricos).

Os modelos cujos resultados mais correspondiam à realidade foram implementados e usados para gerar as cartas de risco, que foram integradas na plataforma do projeto com sucesso e estão disponíveis para visualização.

Foi também disponibilizada uma página com o histórico dos índices de risco para cada parcela e uma outra página mais direcionada à análise histórica do risco de todas as estações disponíveis.

Os objetivos do projeto foram concluídos com sucesso, tendo sido fornecida informação útil aos empresários agrícolas para que possam melhor proteger as suas culturas do brócolo.

5.2 Trabalho futuro

Por falta de recursos financeiros e humanos, só foi possível recolher dados de sintomas de uma parcela na região do Ribatejo, o que poderá ter resultado num sobreajuste dos modelos a esta região. O próximo passo na otimização dos modelos desenvolvidos envolve a instalação de estações meteorológicas locais em parcelas noutras zonas do país.

Deve então ser feita também a observação de sintomas em todas as parcelas e devem ser reanalisados os modelos, procedendo inclusive a alteração de certos parâmetros, caso se justifique, de modo a obter os modelos que melhor se ajustem a todo o território português.

Neste momento, as cartas de risco presentes na plataforma online apresentam apenas o risco para o território português continental.

Os dados de risco para o arquipélago dos Açores e para a Região Autónoma da Madeira já se encontram na base de dados do projeto, sendo apenas necessário encontrar um shapefile relativo a Portugal que inclua estas duas regiões e substituir o atualmente em utilização.

Um outro trabalho futuro interessante seria expandir o sistema e permitir acolher novas culturas além do brócolo, juntamente com os modelos relativos aos seus inimigos. Assim, um empresário agrícola que tenha várias culturas pode trabalhar apenas com uma plataforma e gerir de forma mais fácil a aplicação de produtos fitofarmacêuticos.

BIBLIOGRAFIA

- [1] N. Ahmed. *Pesticide use in periurban areas*. Vol. 2012. 37. 2012.
- [2] *Alternariose na cultura da couve-brócolo para indústria*.
- [3] M. de Terras de Bouro. *Manual de Agricultura Biológica - Terras de Bouro*. URL: http://www.actuar-acd.org/uploads/5/6/8/7/5687387/manual_ab_terras_bouro.pdf.
- [4] S. Bregaglio. "Definition and implementation of plant disease simulation models in interaction with crop models, aiming at forecasting the impact of climate change scenarios on crop production". Tese de doutoramento. 2011-2012.
- [5] G. Calow. *Delia radicum - Delia radicum*. URL: <http://www.naturespot.org.uk/species/delia-radicum>.
- [6] S. Collier R.H. Finch. "Thermal requirements for cabbage root fly, *Delia radicum*, development." Em: (1988).
- [7] F. Ferreira. *Cartas de Risco para Pragas e Doenças das Plantas*. 2016.
- [8] A. Fraval. *Cabbage fly, Cabbage root fly*. URL: <https://www7.inra.fr/hyppz/RAVAGEUR/6delrad.htm>.
- [9] C. of Grande Prairie No. 1. *Canola Alternaria Black Spot*. URL: <http://www.countygp.ab.ca/EN/main/departments/agriculture/pest-disease-control/alternaria-black-spot.html>.
- [10] C. Hong, B. D. Fitt e S. Welham. "Effects of wetness period and temperature on development of dark pod spot (*Alternaria brassicae*) on oilseed rape (*Brassica napus*)". Em: *Plant Pathology* 45.6 (1996).
- [11] *How to Manage Pests*. URL: <http://ipm.ucanr.edu/WEATHER/ddconcepts.html>.
- [12] Huber, L, Gillespie e TJ. "Modeling leaf wetness in relation to plant disease epidemiology". Em: *Annual review of phytopathology* 30.1 (1992).
- [13] R. Kennedy e A. Graham. "Final report HDC project FV53b". Em: (1994).
- [14] R. Kennedy, K. Phelps e A. J. Turner. "Prediction of Sporulation by *Alternaria Brassicae* and *A. Brassicicola* on *Brassica Napus*". Em: (1999).
- [15] K. Kim, S. Taylor, M. Gleason e K. Koehler. "Model to enhance site-specific estimation of leaf wetness duration". Em: *Plant disease* (2002).

- [16] Kruit, R. J. Wichink, van Pul, W. A. J, Jacobs, A. FG, Heusinkveld e B. G. "COMPARISON BETWEEN FOUR METHODS TO ESTIMATE LEAF WETNESS DURATION CAUSED BY DEW ON GRASSLAND". Em: (2004).
- [17] M. G. Lawrence. "The relationship between relative humidity and the dewpoint temperature in moist air: A simple conversion and applications". Em: *Bulletin of the American Meteorological Society* (2005).
- [18] R. Magarey, T. Sutton e C. Thayer. "A simple generic infection model for foliar fungal plant pathogens". Em: *Phytopathology* 95.1 (2005).
- [19] Martins, Aguiar, Almeida, Ferreira e L. Leite. "A mosca do nabo (*Delia radicum*)". Em: (2005).
- [20] R. McDonald e M. Sears. "Effects of root damage by cabbage maggot, *Delia radicum* (L.)(Diptera: Anthomyiidae), on yield of canola, *Brassica campestris* L., under laboratory conditions". Em: *The Canadian Entomologist* (1991).
- [21] J. Monteith. "Dew". Em: *Quarterly Journal of the Royal Meteorological Society* 83.357 (1957).
- [22] M. Murray. "Using degree days to time treatments for insect pests". Em: (2008).
- [23] J. J. Prenger e P. P. Ling. "Greenhouse Condensation Control Understanding and Using Vapor Pressure Deficit (VPD)". Em: (2001).
- [24] SciPy. URL: <https://www.scipy.org>.
- [25] P. Verma, G. Saharan et al. "Saskatoon Research Station Technical Bulletin 1994-6E. Monograph on Alternaria diseases of crucifers." Em: *Saskatoon Research Station Technical Bulletin 1994-6E. Monograph on Alternaria diseases of crucifers*. (1994).
- [26] X. Yin, M. J. Kropff, G. McLaren e R. M. Visperas. "A nonlinear model for crop development as a function of temperature". Em: *Agricultural and Forest Meteorology* 77.1 (1995), pp. 1–16.